CHARACTERIZING PATTERNS OF ROUND GOBY (*NEOGOBIUS MELANOSTOMUS*) DISTRIBUTION USING MORPHOMETRICS, OCCUPANCY MODELING, AND POPULATION GENETICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mohd Zafri Hassan

May 2013

CHARACTERIZING PATTERNS OF ROUND GOBY (*NEOGOBIUS MELANOSTOMUS*) DISTRIBUTION USING MORPHOMETRICS, OCCUPANCY MODELING, AND POPULATION GENETICS

Mohd Zafri Hassan, Ph. D.

Cornell University 2013

In addition to anthropogenic mechanisms of dispersal, establishment of round goby, *Neogobius melanostomus*, in the Laurentian Great Lakes has generally benefitted from continuous connections among the lakes and accessible shorelines. This dissertation (1) examined the effects of distance in affecting round goby invasion and the native pumpkinseed sunfish *Lepomis gibbosus* population along the lake shoreline as reflected by morphological differences, (2) estimated the co-occupancy of round goby and native benthic fish (i.e. darters) in the tributary system and (3) analyzed the impact on population genetics resulting from round goby invasion in streams as inferred from gene flow. The study area spanned the Lake Ontario shoreline and its tributaries in the northwestern area of New York. Morphological variation was found to be significantly different among sampled sites for pumpkinseed, but was not found to be significant for round goby. For pumpkinseed, samples from river-mouth sites showed shallower body depth than that of samples from bay-area sites. The degree of morphological variation for pumpkinseed increased with distance but the effect was not significant for round goby. The study further showed when estimating species occurrence using presence/absence data, models that assumed perfect detection ($p = 1$) when round goby are present can be underestimated as much as 5 to 60% as compared to models

that assumed imperfect detection ($p < 1$) when the probability of occupancy of darters was considered. While the occupancy of streams by round goby decreased with link magnitude (a proxy for stream size), the probability of occupancy by round goby can also be underestimated by 2 to 15% when imperfect detection is not accounted for. Lastly, I found population genetic differentiation of round goby at fine scales, where individuals collected in the invaded areas exhibited at least two distinct lineage clusters that predictably coincided with the closest major waterways, Lake Ontario and the Erie Canal. Also the round goby population did not exhibit signs of founder effects with respect to stream network structure in the study area. In conclusion, identification of the differential response in body-shape morphology suggests that localized threat on limited-dispersal species like pumpkinseed by invasive generalist species like round goby could affect selection. Where habitat heterogeneity contributes little to no impact on round goby colonization, estimation of colonization and impacts on native benthic fish in the invaded areas should account for imperfect detection of sampling. Information from this study can be adapted to develop early invasive species detection in developing control management of further invasion and for conservation of native species.

BIOGRAPHICAL SKETCH

Mohd Zafri Hassan hails from Malaysia. He received a B.Sc. concentrating in fishery science from University Putra Malaysia (UPM) in 2002. Upon graduation, he worked as a research assistant at UPM working on phyto-chemistry before joining the university's campus in Borneo as a tutor. While there, he taught fish biology and aquaculture-related courses. He then went to the graduate school at UPM, earning a M.Sc. in Aquaculture Technology in 2007 where he studied a native catfish, *Pangasius nasutus*, reproductive biology in the longest river in Peninsular Malaysia. He received a scholarship from the Ministry of Higher Education of Malaysia to pursue a doctoral degree at Cornell University in 2008. He initially worked in the Department of Natural Resources, Cornell University under Dr. Mark Bain's supervision and then completed the doctoral program under Dr. Patrick J. Sullivan. He returns to his home country to work at University Putra Malaysia and looks forward to doing research in the rich faunal diversity of Borneo.

ACKNOWLEDGEMENT

I would like to express the deepest appreciation to my late committee chair, Professor Mark B. Bain (1952-2012), who had the attitude and the substance of a great educator in which he continually and convincingly conveyed a spirit of adventure in regard to research and scholarship. Without his trust, pursuing a doctoral program at Cornell University would not have been possible to begin with.

I would like to thank Dr. Patrick J. Sullivan who took the role as my dissertation advisor upon Dr. Bain's passing and for his persistent help in sharing ideas and suggestions that guided me through, especially during the critical time of the data analyses and writing processes. Many thanks are also due to Dr. Amy R. McCune and Dr. William B. Fisher who served as my dissertation committee members and provided helpful insights into developing research questions and writing the dissertation.

The last chapter of my dissertation owed very much to Dr. Matthew Hare who allowed me to work in his genetic laboratory. His suggestions in building up my research questions on round goby's population genetics and comments on the chapter manuscript are greatly appreciated. Harmony Borchardt-Wier, who works in his laboratory, has been really helpful in lending her hand and providing useful advice from the beginning until the final stages of the genetic laboratory work.

I fervently thank the New York State Department of Environmental Conservation for sampling permits and the private land-owners for their permission to access their

property in the name of science. I would sincerely like to thank Gregory Anderson who introduced me to the species occupancy modeling and untiringly helped me in the field during sample and data collection for two consecutive years of 2010 and 2011. His excellent knowledge and skills, in biology and ecological models, gain my admiration. My brilliant and insightful friend, Morgan L. Ruelle, who has not only been a very good friend, but also patiently helped me with his editorial and critical comments on my morphometric and genetic chapters. Also, I wish to thank countless of people that I have not met but I have asked questions of during the data analyses processes.

The doctoral program allowed me to get to know a great people that made the academic pursuit more special, especially: Claire Southamher, Kimberly Falbo, Zarirah and Dzarifah Zulperi, Murni Karim, Fadzli Fuzi, Wan Lutfi Wan Johari, Tan Teng Ju, Nor Afizah Mustapah, Adi Md Sikin, Chuan Liao, Alex Kudrestav and Wang Sihai.

Many thanks are also due to the Government of Malaysia through the Ministry of Higher Education and University Putra Malaysia for providing financial support throughout the study period.

Last but not least, I would like to tremendously thank my mother and father who continuously provide inspiration through their hardship in raising their thirteen children. Their love and support throughout my education has greatly taught me the lessons of life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
CV	Canonical variate
CVA	Canonical variate analysis
CS	Centroid size
°C	Celsius
DIST	Distance
DNA	Deoxyribonucleic acid
ESRI	Environmental Systems Research Institute
GIS	Geographical information system
GPA	Generalized Procrustes analysis
HUC	Hydrological Unit Code
LMAG	Link magnitude
HWE	Hardy-Weinberd equilibrium
MANCOVA	Multivariate analysis of covariance
MANOVA	Multivariate analysis of variance
MCMC	Markov chain Monte Carlo
MS-222	Tricane mesylate
PCR	Polymerase chain reaction
NHD	National Hydrologic Database
s.e	Standard error
s.d	Standard deviation

SL	Standard length
SUBS	Substrate
TPS	Thin plate spline
USGS	United State Geological Survey
V	Voltage
SIF	Species Interaction Factor
TPM	Two-phase mutation model
ns	Not significant

LIST OF SYMBOLS

Δ	Delta
α	Alpha
cm	Centimeter
km	Kilometer
mm	Millimeter
m	Meter
p	Probability of detection
ψ	Psi
w_i	Akaike information criterion weight of evidence
mg	Milligram
pH	Hydrogen ion concentration
µL	Micro-liter
H_2O	Water
F	F-statistics
K	Number of population structure
N_A	Number of alleles
h	Gene diversity
LnP	Posterior log-likelihood

PREFACE

Species introduction outside their native range becomes a global phenomenon with increased international trade (Rixon et al. 2005; Westphal et al. 2007). Although not all introduced species become established in the new recipient environment, species that do, often have been reported posing economic losses, altering the composition of communities and reducing ecosystem services when they spread in the introduced range, and are therefore regarded as invasive (Mills et al. 2004, Pimentel et al. 2005, Cucherousset and Olden 2011, Dettmers et al. 2012, but see Keller et al. 2007, Gozlan 2008).

The Laurentian Great Lakes region is one of the prominent impacted areas of species invasion in the world, where the invasive species were primarily released into the region by accidental ballast-water transfer of transoceanic shipping (Grigorovich et al. 2003; Holeck et al. 2004). One of the invasive species that is currently exhibiting negative consequences in the region is round goby, *Neogobius melanostomus*. Like many introduced species in the region, round goby, was accidentally introduced through ballast water transfer from its native range of Black, Caspian and Azov seas and the lower reaches of associated rivers (Marsden et al. 1996, Stepien et al. 2005, Brown and Stepien 2009). Since its first sighting in the early 1990's in St. Clair river, Ontario (Jude et al. 1992) the round goby population has increasingly expanded across the Great Lakes and the tributary systems (Clapp et al. 2001, Phillips et al. 2003, Krakowiak and Pennuto 2008, Poos et al. 2009, Kornis et al. 2010). The establishment

of round goby in the region has benefitted from continuous connections among the Great Lakes and through accessible shorelines for both natural and anthropogenic dispersal.

Following round goby invasion in the streams, the majority of studies have shown adverse effects caused by goby through direct and indirect predation and competition that it may and have caused on the faunal communities in the invaded range, although some studies have reported highly abundant round goby provided more food resources for some predatory fishes (Johnson et al. 2005; Madenjian et al. 2011). In addition to shipping activities within the region that usually facilitate translocation (LaRue et al. 2011), rapid secondary invasion of round goby into the tributaries is also attributed to other human activities such as transfer via bait-buckets (Marsden et al. 1996; Vander Zanden and Olden 2008). Evidently, much of the ability of rapid secondary invasion of round goby is attributed to its natural short- and long-distance dispersal mechanisms coupled with among others, its aggressive behavior, its ability to reproduce quickly in high numbers, its tolerance to most habitats, its broad diet preferences and its lack of a larval stage that expedites population turnaround (Ray and Corkum 2001, Simonovicå et al. 2001, French III and Jude 2001, Taraborelli et al. 2009, Bronnenhuber et al. 2011).

Whereas the biological attributes of successful establishment are generally understood, knowledge gaps in the environmental factors that facilitate round goby range expansion remain a challenge as the species continues to invade new environments.

Greater knowledge in this area will provide information that can help mitigate their impacts. As research on the species continues, habitat and diet preferences by round goby have provided indications that: (1) There will be areas that are more susceptible to invasion than others; (2) Local-scale ecological characteristics can constrain or at least mitigate impacts on native recipient communities. My dissertation hopes to determine whether those indications are actually reflected in nature, by addressing three research objectives that are presented in the three research chapters in this dissertation. The objectives are:

1) To examine the effects of distance in influencing round goby invasion along the lake shoreline as reflected by morphological differences
2) To estimate the co-occupancy of round goby and native benthic fish (i.e. darters) in the tributary system
3) To infer the population dynamics influencing round goby invasion in streams inferred from gene flow

In Chapter 1, I explored the roles of distance among habitat patches along the Lake Ontario shoreline of New York in influencing body shape variations in round goby. Round goby typically expand into the tributaries from the mouth of rivers of open-bay shoreline where productivity is lower than in wetlands (Cooper et al. 2007). This is because wetlands are typically higher in productivity, attracting more diverse species that consequently provide resistance to invasion (Coulter et al. 2012). If open-water-and-wetlands occur between each other along the lake shoreline, the interspersed near-shore lake patterns may consequently affect the mechanisms round goby to invade and

persist in the tributaries.

Diversification of a species along ecological space can result in variation in morphology, life history or behavior (Sanford and Worth 2010). Of the three types of diversification, morphological differentiation is most easily detectable, and its detailed assessment is important in understanding the degree of local adaptation in response to changes in a species' environment or niche. In this chapter, I analyzed the body shape of fish because shifts in such a trait have been related to swimming performance and feeding (Haas et al. 2010; Langerhans 2008; Vila-Gispert et al. 2007). Body shape can also influence fitness by affecting foraging success, fecundity and predator avoidance (Langerhans and Reznick 2010). I specifically addressed two questions: (1) Does body morphology differs between locations and if so what is the direction of morphological variation; (2) Does connectivity between habitat patches along the shoreline predict the level of variation?

Using location and geographic distance as a surrogate for testing morphological variation revealed that round goby and pumpkinseed sunfish achieved solution to sustain populations in different ways. The extent to which morphological differences represent alternative adaptation for pumpkinseed needs further study. For round goby, non-significant body shape variation among locations in our study is perhaps a character of a successful invader (L'avrincikova et al. 2005). It is also possible that selective pressures are resulting in different phenotypes on round goby, as recent studies indicate genetic diversification of round gobies in the Great Lakes region

(Bronnenhuber et al. 2011; Nolte 2011). Coincidently, pumpkinseed is considered as an invasive species in Europe. Studies between native and invasive pumpkinseed populations inhabiting fluvial and lacustrine waterbodies showed that pumpkinseed acquire adaptive external morphologies in the invaded range related to locomotion such as median-fin size and placement and also body width through successive generations (Yavno et al. 2012). Such adaptation processes might have altered the species genetically or phenotypically, therefore, one can expect this to be easily observed. I further suggest experimental studies to be carried out to examine the effect of divergent selection so as to further quantify the potential evolutionary effects of native and invasive species in the Laurentian Great Lakes region.

Chapter 2 uses species occupancy models to accounting for false detection, which is critical for estimating the presence or absence of a species. False absence or failure to detect a species at a site when it is actually present is expected to occur in natural settings especially where populations are in low abundance and elusive (Mackenzie et al. 2004; MacKenzie et al. 2002). The chapter follows an extension of occupancy modeling where the presence of an inferior species is assumed to be conditional on the presence of superior species (Richmond et al. 2010).

For this chapter, I explored the interaction between *N. melanostomus* and multiple darter species in stream tributaries of New York State that drain into Lake Ontario. I was interested in understanding the roles of habitat criteria in facilitating the distribution of these benthic species while accounting for heterogeneity in detection

and variable stream characteristics. In developing models for co-occurrence patterns of round goby and darter species, I predict that: (1) round goby should have a higher probability of occupancy in areas closer to Lake Ontario, the presumed invasion 'entrance' to tributary systems, than darter because round goby invasion could potentially lead to population displacement of darter as shown in other studies; (2) in small streams where resources are presumably more limited, darter occupancy should be lower when round gobies are present because interference competition should increase with co-occupancy; (3) the probability of occupancy for both species should increase with stream size because larger streams should provide a wider range of microhabitats and contain more resources to support both species; (4) Round gobies are more aggressive and in high abundance when present than darter. Therefore competition for space and resources would decrease the probability of detection of darter.

A total of 51 species from 13 families were detected during the sampling period. Besides round goby, overall I found seven darter species namely: greenside darter *Etheostoma blennoides,* rainbow darter *E. caeruleum*, Iowa darter *E. exile*, fantail darter *E. flebellare*, tessalated darter *E. olmstedi*, logperch *Percina caprodes* and blackside darter *P. maculata* in the sampling area. For both round goby and darters, the best-supported detection model was the model that included stream link magnitude as a detection covariate, where darter detection probabilities appeared to be conditional on round goby detection when both species were present. Models that assumed the occupancy of darter was conditional on round goby received stronger

support as compared to models that assumed no relationship between the two species groups. The top model, which accounted for imperfect detection, provided much better description of the data than the model that assumed detection probability equal to 1. I found that larger streams are more susceptible to round goby invasion despite the fact that some sites of smaller streams were closer to Lake Ontario, the presumed invasion source. The results indicated round goby is expanding its range into upstream tributary systems and show an overlapping range with darter. Both round goby and darter were more likely to co-occupy smaller streams (lower link magnitude) than sites of larger magnitude. I found that species interaction, aggregation or competition between round goby and darter decreased as the magnitude of stream covariates increased.

I further explored the roles of stream network in facilitating or impairing the round goby invasion in Chapter 3. In this chapter, I postulate that the geography of the stream network limits dispersal between streams and therefore inhibits gene flow. Round goby populations in the Great Lakes Region are known to descend from multiple genetic sources from their native range in Eurasia as shown by the lack of founder effect in the invaded areas (Brown and Stepien 2008; Dillon and Stepien 2001). Given that natural dispersal of round goby has occurred at sufficient rates to enable its spread throughout the tributary system within a relatively short period of time, and there has been limited time for any genetic differentiation (on the order of two decades), a reasonable null hypothesis is that little to no genetic differences can be observed between goby populations within the study area at this time.

However, previous studies of round goby indicate that the species can achieve high between-site genetic differentiation in less than ten-generations, even within a relatively small spatial range (Björklund and Almqvist 2010; LaRue et al. 2011). In the chapter, I tested two alternative hypotheses that are not mutually exclusive of population differentiation to counter my null hypothesis: (1) if natural dispersal has been infrequent and limited to isolated founding events, new upstream populations will remained isolated from the larger downstream populations, resulting in significant observable genetic differences; (2) competition with native benthic fish may limit upstream population growth and contribute to higher rates of allele fixation through genetic drift.

I analyzed the genetic structure of round goby populations in tributaries that drain into Lake Ontario from northwestern New York. To test the first alternative hypothesis, I measured the relationship between genetic differentiation, the position in the riverscape (i.e. link magnitude), and geographic distance from the major round goby populations in Lake Ontario and the Erie Canal. For the second alternative hypothesis, I examined the relationship between the genetic diversity of round goby and the presence and absence of native darter at each site. This study was conducted at the same sites described in the preceding chapter (Chapter 2), and genetic diversity was analyzed at neutral molecular markers, including eight microsatellite loci within the nuclear genome of round goby.

In the present study we found the round goby populations showed high within-site

genetic diversity and significant local scale genetic differentiations among study sites, corroborating the studies conducted along the Lake Michigan and in the Baltic Sea (Björklund and Almqvist 2010, LaRue et al. 2011)**.** Apart from Lake Ontario, we also found the goby populations also invade northwest New York tributaries from the Erie Canal as revealed by a distinct population structure that are coincided with the location of our study sites as well as genetic admixture between those two major watercourses. However, we found lack evidence that stream networks can constrain population mixing. Although the mechanisms for genetic unification and differentiation include short-distance dispersal to nearby sites and longer-distance dispersal to remote sites (Bronnenhuber et al. 2011; LaRue et al. 2011), along tributaries, dams and waterfalls may limit dispersal of round goby upstream (Kornis and Vander Zanden 2010). Our results revealed evidence of limitations to the natural dispersal of round goby as shown by significant pairwise F_{ST} values among many pairs of populations, admixture in Bayesian assignment analysis, and spatial patterns of genetic differentiation. Pattern of genetic differentiation (based on F_{ST}) between sites indicate that the distances between sites are not the only factors determining differences between local populations of round goby in northwestern New York. F_{ST} for some site pairs of adjacent distance were more significantly different that site pairs of farther distance.

However, we found lack evidence that stream networks can constrain population mixing. Although the mechanisms for genetic unification and differentiation include short-distance dispersal to nearby sites and longer-distance dispersal to remote sites (Bronnenhuber et al. 2011; LaRue et al. 2011), along tributaries, dams and waterfalls

may limit dispersal of round goby upstream (Kornis and Vander Zanden 2010). These short- and longer-distance dispersal classes may be caused by different mechanisms; short-distance dispersal may be natural, while long-distance dispersal may be anthropogenic (e.g., commercial shipping or bait bucket transfers) (Bronnenhuber et al. 2011; Hensler and Jude 2007; LaRue et al. 2011; Lynch and Mensinger 2012).

Conclusions:

(1) If given enough generational time and constrained population mixing, round goby is expected to show morphological variation, similar to what has been found in pumpinkseed along the Lake Ontario shoreline, although the characters of variation may differ considering their different habitat niche i.e. benthic (round goby) and pelagic (pumpkinseed).

(2) Round goby prefers larger stream sizes that are hypothesized to provide more resources, whereas occurrences of native benthic fish (Etheostomes and Percinas) were found to decrease. Interpretation of competitive exclusion of round goby on native benthic fish in this regard was not clear, however, because the benthic fish species distribute allopatrically across the study area.

(3) Genetic differences of round goby in the study area was primarily attributed to multiple sources of invasion as evidenced by the presence of two genetic clusters that can be associated with Lake Ontario and the Erie Canal populations. Lack evidence of the role of stream networks in facilitating or constraining upstream areas from the presumable source of introduction that could unify the genetic makeup, and little evidence of admixture individuals at given sites alternatively suggests the invasion in the area might be recent.

(4) Unless the native species is known to distribute equally the same with round goby, effects of round goby invasion on native fish at large spatial scales should be examined based on a single-species approach to avoid over- or under-interpretation of sampling results.

CHAPTER 1

MORPHOLOGICAL ANALYSIS OF AN INVASIVE FISH, ROUND GOBY *Neogobius melanostomus* AND A NATIVE SPECIES, PUMPKINSEED SUNFISH *Lepomis gibbosus* ALONG LAKE ONTARIO SHORELINE

Abstract

Near-shore environments can generate ecological variation that can create differential evolutionary response. I used geometric morphometric analysis to examine these responses in an invasive fish, round goby *Neogobius melanostmous,* and a native fish, pumpkinseed sunfish *Lepomis gibbosus,* along the Lake Ontario shoreline of New York. Specifically, I analyzed the direction of morphological variation and determined whether connectivity between habitat patches along the shoreline predict the level of variation. The directions of morphological variation occur in the same direction for both species although only pumpkinseed showed significant morphological differences across study sites. In round goby, the morphological changes were subtle, occurred mainly in the caudal peduncle region. For pumpkinseed sunfish, the major morphological differences were longer distances between snout and the base of dorsal fin and shorter caudal peduncle in samples collected from three farthest eastern locations than the rest of the sampled locations. Pumpkinseed sunfish showed a significant inverse relationship between morphological differences and geographic distance. Based on the results, significant morphological variation would potentially occur in round goby in the future if new invading individuals do not interrupt

restricted population mixing such that seen in pumpkinseed population in the area. We revealed that geographic distance serves as an important constraint on adaptive diversification, where the level of variation can be constrained by the biological background of the species under study.

1.1 Introduction

Shoreline environments can generate variation in the distribution of nutrients, phytoplankton and larva over spatial scales of thousands kilometers (Meixler et al. 2005; Meixler and Bain 2010; Singkran and Meixler 2008). Whereas the ecological consequences of the near-shore environments on fish community is increasingly well understood (Hinch et al. 1994; Randall et al. 1996; Wei et al. 2004), the investigation of how these persistent spatial differences along the shoreline can impose evolutionary variation within a species is largely unexamined. The lack of research in this area may be because the shoreline is viewed as providing a continuum in which fish are assumed to be genetically similar across especially small spatial ranges. Heterogeneous shoreline environments, in providing varying magnitude of available resources, can promote phenotypic diversification (Sanford and Worth 2010).

However, phenotypic diversification is usually constrained by population mixing despite the heterogeneous shoreline environments (Clarke et al. 2010), in which the degree of mixing may vary interspecifically. This is because different species achieve varying degrees of population mixing, based on their life-history traits, physiology and dispersal strategies (Smith and Donoghue 2008; Wagner and McCune 2009). For

example, where various habitat types intersperse along the Lake Tanganyika shoreline, two trophically specialized cichlids exhibit much more restricted population mixing over sandy habitat than the other trophically opportunistic cichlid (Wagner and McCune 2009).

Thus, understanding general patterns and causes of diversification requires an examination across multiple species and an evaluation of potential constraints on divergence. This is particularly important in the context of climate change and as native fish populations are threatened by introduction of exotic species (García-Berthou 2007; Lee 2002; Olden et al. 2006; Rahel and Olden 2008). Under variation of environmental conditions, external morphological characteristics may arise that may make exotic species better able to address the challenges of a new regime as would be expected to occur via generational selection or phenotypic plasticity (Yavno et al. 2012). Of the three types of diversification i.e. morphology, life history or behavior (Sanford and Worth 2010), morphological differentiation is most easily detectable, and a detailed assessment is important in understanding the degree of local adaptation in response to changes in a species' environment or niche.

We compared body morphology of two distinct fish species: the exotic round goby, *Neogobius melanostomus* (Pallas 1814), and the native pumpkinseed sunfish, *Lepomis gibbosus* (L. 1758), of Lake Ontario. We analyzed the body shape of fish because shifts in such a trait have been related to swimming performance and feeding (Haas et al. 2010; Langerhans 2008; Vila-Gispert et al. 2007). Body shape can also influence

fitness by affecting foraging success, fecundity and predator avoidance (Langerhans and Reznick 2010). We sought to examine whether the Lake Ontario shoreline facilitates or constrains morphological differences in the round goby populations. Since its first reported population establishment in 1990 (Jude et al. 1992), the species has caused adverse ecological impacts across the invaded areas (Pennuto et al. 2010; Phillips et al. 2003; Poos et al. 2009; Walsh et al. 2007). Instead of using evolutionarily similar benthic morphology, we examined the same ecological effect on pumpkinseed sunfish morphology since this species is abundant in the shoreline, and therefore available for comparison. Furthermore, there were multiple ecological and evolutionary biology studies of the pumpkinseed in the same ecological region (Brinsmead and Fox 2002; Gillespie and Fox 2003) that could facilitate interpretation of morphological shifts observed in other fish.

For the present study we addressed two questions: (1) Does body morphology differ between locations and if so what is the direction of morphological variation; (2) Does connectivity between habitat patches along the shoreline predict the level of variation?

1.2 Materials and methods

1.2.1 Sample collection

We collected round goby and pumpkinseed sunfish from ten sites along the Lake Ontario shoreline of New York. Eight of these locations were used for further analysis, based on the criterion that sample sizes were greater than the number of landmark points needed for morphometric analysis (Figure 1.1). Pairwise geographic distance in

kilometers (km) between localities ranged between 17 – 349 km.

All samples included in the following analysis were collected between August 11 and September 2, 2012, in deep littoral areas, including four bays and four river outlets (Table 1.1). Sampling was conducted using an electrofishing boat for 1.5 hours of sampling effort in each location. The boat was equipped with a Smith-Root Type VI-A transformer, a Honda 5000 watt generator, and two anode probes suspended in front of the boat, and the boat hull served as the cathode array. Within each site, sampling effort was equally distributed across each available microhabitat within approximately 500m radius of the boat launching area. Electrofishing was conducted by placing 25 to 250 V of direct current into the water for about 2-3 minutes for each microhabitat. Round gobies were collected around rocky (boulders, cobbles) or soft substrate (silt, sandy) areas near the water's edge. Pumpkinseed sunfishes were collected in the adjacent deeper areas with submerged woody vegetation. Stunned fish appearing between the electrodes were retrieved using dip nets. Collected samples were transferred into a water bucket on board the boat, euthanized with MS-222 and sorted. All samples were stored at -5°C in a portable freezer and transferred to -20°C freezer after the conclusion of fieldwork prior to analysis.

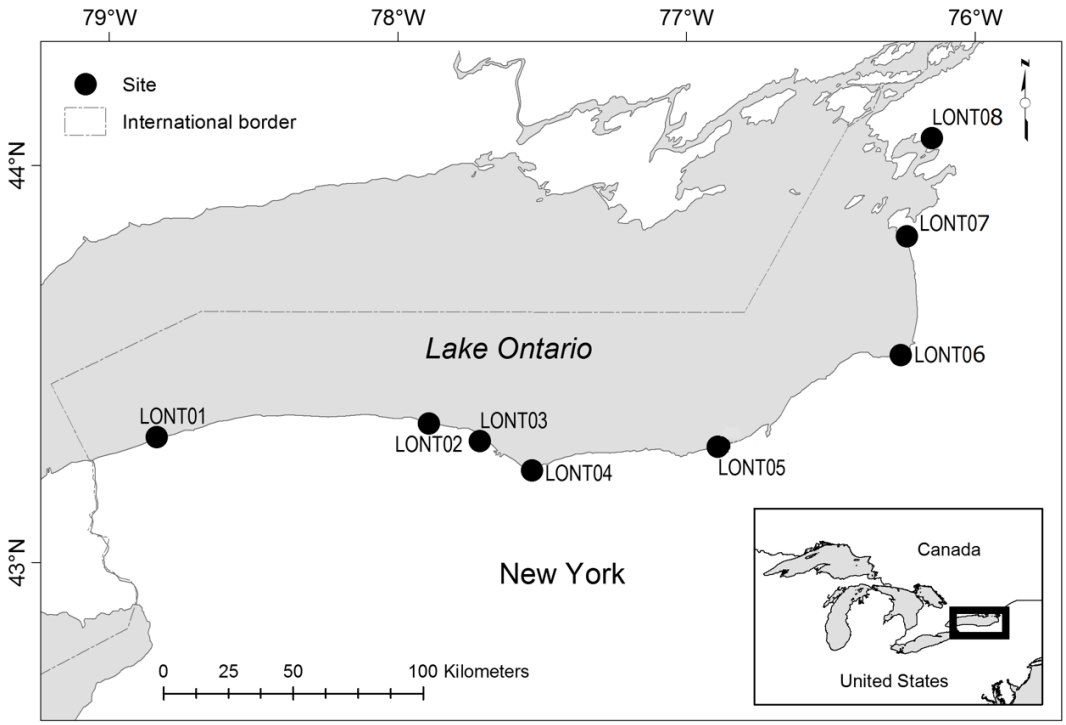


Figure 1.1. Map showing the sampling location along the Lake Ontario shoreline of New York. LONT01 = Tuscarora Bay (Niagara county), LONT02 = Sandy Creek (Monroe county), LONT03 = Braddock Bay (Monroe county), LONT04 = Irondequoit Bay (Monroe county), LONT05 = East Bay (Wayne county), LONT06 = Little Salmon River (Oswego county), LONT07 = Stony Creek (Jefferson county), LONT08 = Chaumont River (Jefferson county).

Table 1.1. Geographic characteristics of the sampled localities.

Site code	Water body (County)	Latitude	Longitude	Habitat type	Nearby land-use cover (>50%)
LONT01	Tuscarora Bay (Niagara)	43.3173	-78.8362	Bay	Developed, forest
LONT02	Sandy Creek (Monroe)	43.3514	-77.8918	River mouth	Crops
LONT03	Braddock Bay (Monroe)	43.3068	-77.7160	Bay	Wetlands
LONT04	Irondequoit Bay (Monroe)	43.2330	-77.5349	Bay	Developed
LONT05	East Bay (Wayne)	43.2928	-76.8910	Bay	Forest
LONT06	Little Salmon River (Oswego)	43.5242	-76.2561	River mouth	Wetlands, shrub
LONT07	Stony Creek (Jefferson)	43.8222	-76.2349	River mouth	Wetlands, crop
LONT08	Chaumont River (Jefferson)	44.0679	-76.1485	River mouth	Pasture/hay

1.2.2 Image acquisition

Each specimen was thawed and placed on its right side on a platform overlain with a gridded rubber-mat that served as a scale for analysis of digital images. Specimens were placed so as to attain a normal lateral-view posture. Excess water was patted dry to reduce reflection from the body. A digital photograph of each specimen was taken using a Fujifilm FinePix F610 6.0 megapixels camera. The camera was attached to a copy stand so that the distance between the specimen and the lens is standardized for every image. In order to assist with landmark digitization, laboratory needles were placed on each specimen's landmark points to identify homologous landmarks.

Fish were sexed by visual observation of the urogenital papilla, which is pointed in males and more broad and oval in females of both round goby and pumpkinseed. The presence of testes and ovaries was checked by dissecting a sub-sample of random selected individuals to ensure the accuracy of sexing by external visual inspection.

1.2.3 Landmark digitization and superimposition

Photographed images were converted to thin plate spline (TPS) format in tpsUtil, http://life.bio.sunysb.edu/morph/. Ten homologous points and insertions of the round gobies and pumpkinseed sunfishes were digitized in tpsDig2 version 2.15 to generate Cartesian coordinates. Digitization for all images was recorded twice. The ten homologous points and insertions on the external body morphology were: 1. tip of the upper jaw; 2. insertion of the first spine of the dorsal fin; 3. insertion of the last ray of the dorsal fin; 4. upper base of the caudal peduncle; 5. lower base of the caudal peduncle; 6. insertion of the last spine/ray of the anal fin; 7. insertion of the first spine/ray of the anal fin; 8. projected point that is on the same plane of the base of the pelvic fins; 9. upper base of the pectoral fin; and 10. center of the left eye (Figures 1.2 and 1.3).

All specimen images were then superimposed and aligned by performing a landmark-based method and a generalized Procrustes analysis (GPA), a feature available in MorphoJ (Klingenberg 2011). The superimposition method standardizes each Cartesian unit to centroid size and minimizes differences in translation and rotation of all specimens using a least-squares algorithm. The superimposition method projects

shape coordinates into a Euclidean space tangent to the Procrustes shape space producing Procrustes coordinates that removes the information not related to shape. The Procrustes superimposition translates the centroid of the shapes to (0,0); the *x* coordinate of the centroid is the average of the *x* coordinates of the landmarks of an individual, and the *y* coordinate of the centroid is the average of the *y*-coordinates. Shapes are scaled to unit centroid size, which is the square root of the summed squared distances of each landmark to the centroid. The configuration is rotated to minimize the deviation between the centroid size and a reference, typically the mean shape. In geometric morphometric, centroid size is used as a measure of size and typically inversely related to the length measurement e.g. standard length (Zelditch et al. 2004).

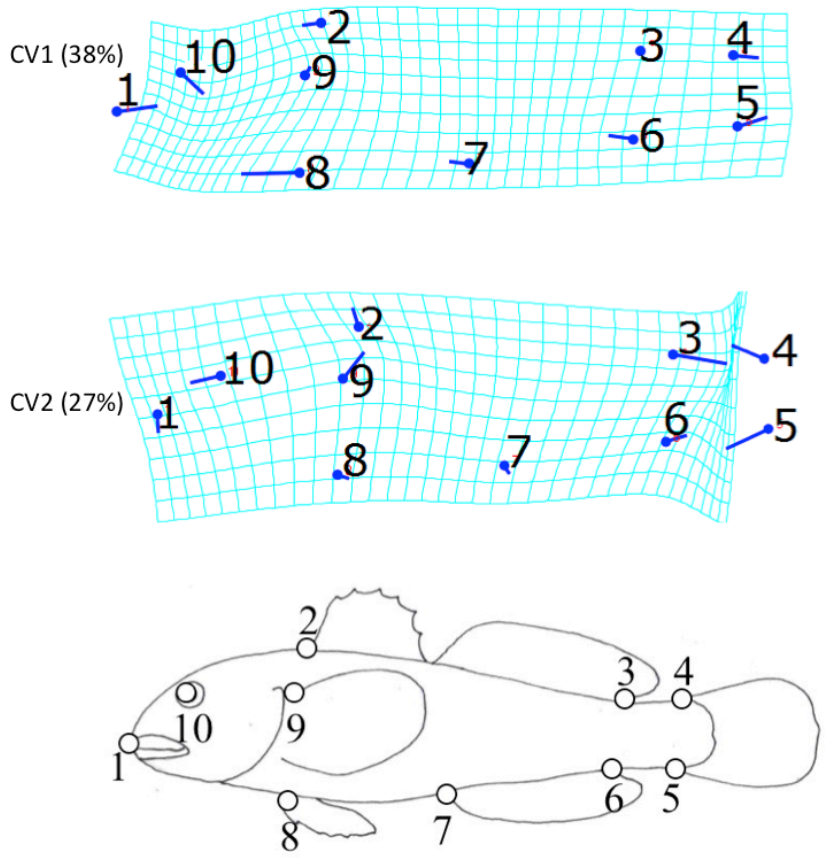


Figure 1.2. Transformation grid showing corresponding shape changes of the canonical variate scores (CV) of round goby *Neogenobius melanostomus* for present study with corresponding percentage of variation in parentheses. CV1 accounts for 38% variation in specimen size, while CV2 accounts for 27% variation in caudal peduncle length. The lollipop symbols in the transformation grid are pointing the direction of shape changes across samples. Tested homologous landmarks are shown at the bottom of the figure (see Materials and Methods section for explanation of the landmarks point numbers).

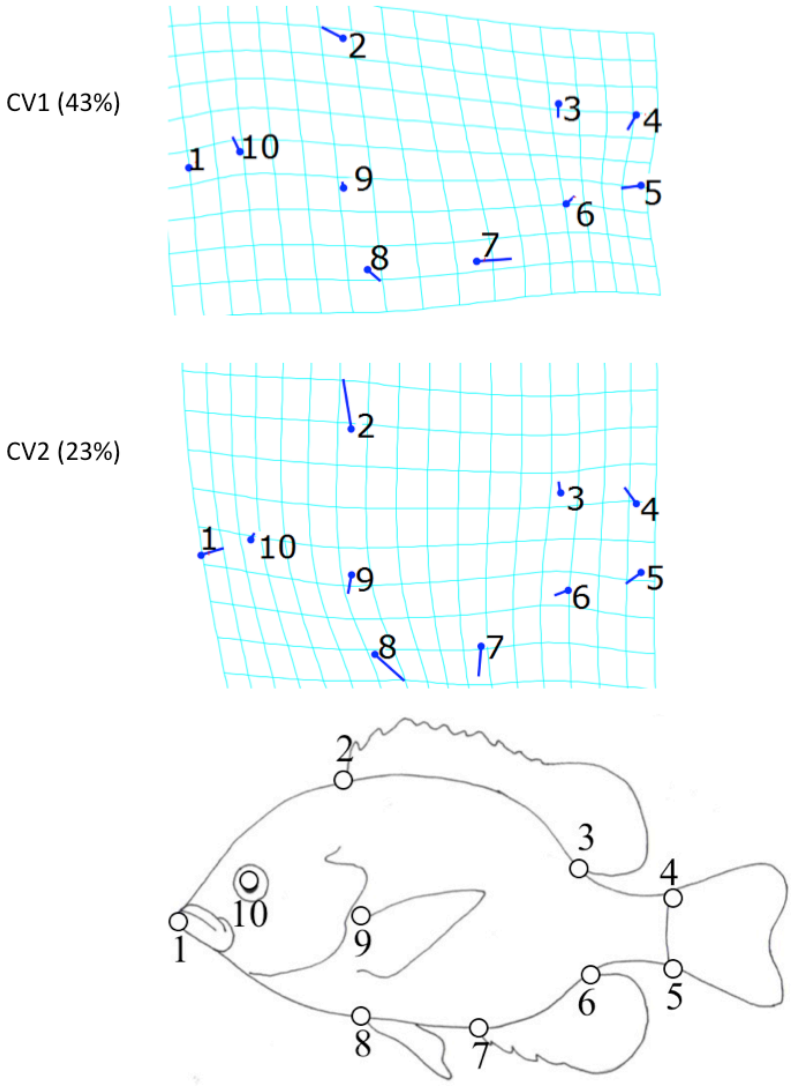


Figure 1.3. Transformation grid showing corresponding shape changes of the canonical variate scores (CV) of pumpkinseed *L. gibbosus* for present study with corresponding percentage of variation in parentheses. CV1 accounts for 43% variation in the head region and caudal peduncle, while CV2 accounts for 23% variation in body depth. The lollipop symbols in the transformation grid are pointing the direction of shape changes across samples. Tested homologous landmarks are shown at the bottom of the figure (see Materials and Methods section for explanation of the landmarks point numbers).

1.2.4 Preliminary statistical procedures: size-correction after controlling for allometric effects

Analysis of morphometric characteristics can be confounded by allometric growth i.e. the larger the size, the larger the ontogenetic shape changes may be. Statistically, allometric growth can be identified by as significant slope in the regression when shape is plotted against size and indicating that a size-correction was needed for the shape data, which is done by using the residuals of the regression model. The residuals are essentially a 'size-corrected' shape.

Shape variables from each sample were regressed on centroid size. A significance test was performed separately in order to determine whether there was any location (site) with a significant regression slope for either of the study species. Significant regression slopes in at least one location would require a standardized size correction for all specimens (Viscosi and Cardini 2011). We further tested whether a single common regression slope could be used to explain the shape model, before using the residuals from the common slope as the 'size-corrected' shape. Multivariate analysis of covariance (MANCOVA) was performed on the Procrustes coordinates per species in order to determine whether the mean shape across localities had similar trajectories by taking into account the size of the specimens. The following effects were included in the MANCOVA test: 1. covariate – size was served as the covariate to control the effect of size; 2. main effects - simultaneously fit effect-specific multivariate linear regressions with each effect having its own slope – besides location, sex of the species was served as another main effect because sexual dimorphism would potentially be a

source of intraspecific variation; 3. interactions – all main effects and covariate interactions were included so that the effect-specific lines were forced to be parallel. Lack of significance of this interaction would mean that the slopes are effectively the same after accounting for all other effects. Regression of all individuals pooled by locality on SL was performed when the MANCOVA interaction was not significant. The second regression was repeated with 1000 permutations to test for slope significance. Residuals from the second regression, all 'size-corrected' shapes were used for all subsequent analyses.

1.2.5 Canonical variate analysis of shape differences between localities

Differences between localities were analyzed by Canonical Variate Analysis (CVA), which provides an ordination that maximizes the separation of the locality means relative to the variation within locality. CVA simultaneously uses all groups in a dataset to compute axes of maximal group separation relative to a shared estimate of within-group variation. The CVA analysis was performed in MorphoJ version 1.03b (Klingenberg 2011). The MANOVA Hotelling's pairwise comparison of the localities with a Bonferoni correction and squared Mahalanobis distance and cross-validation with jackknifing was performed in PAST version 2.15 (http://folk.uio.no/ohammer/past/). Mahalanobis distance measures the variance and correlation of *x* and *y* variables for two individuals between two localities i.e. shape differences relative to the variation in the total sample. Small Mahalanobis distance between samples indicates they are more similar than that of sample with larger Mahalanobis distance.

1.2.6 Visualization of shape changes

A transformation grid, a feature in MorphoJ was used to illustrate shape changes associated with canonical variate (CV) scores obtained from CVA. The CVA illustrates the increased and decreased variation for a particular body region among locations, as shown by compressed or expanded grids. A lollipop graph was overlaid on the transformation grid to aid the interpretation of the CV's change by showing the direction of the shifts of landmark positions (Figures 1.2 and 1.3). Each line starts with a dot at the location of the landmark in the starting (mean) shape. The length and direction of the line indicates movement of the respective landmark from the starting shape to the target shape (e.g. the mean shape plus the shape change that corresponds to an increase of 15 units of Procrustes distance in the direction of the CV1). The transformation grid shows the shape change as a deformation of a rectangular grid using the thin-plate spline.

1.2.7 Testing between-locality differences on geographic distance

Spatially limited population mixing is commonly observed phenomenon in natural populations, that would result in population genetic pattern known as 'isolation by distance' (Wright 1943) and consequently increased in phenotypic divergence (Hendry et al. 2001). Pairwise geographic distance in kilometers (km) between localities was measured in ArcGIS, by measuring the distance between sites within a 500 meter buffer from the shoreline (the shortest possible distance traveled by the study species). We tested the relationship between-locality morphological differences and geographic distance by regressing the Mahalanobis distance for each site-pair (obtained from the

CV analysis) against the geographic distance between those sites (obtained from GIS analysis). The examination of the relationship was performed for each species respectively.

1.3 Results

For clarity, we coded our sample localities according to their geographic position from west to east; LONT01 for Tuscarora Bay, LONT02 for Sandy Creek, LONT03 for Braddock Bay, LONT04 for Irondequoit Bay, LONT05 for East Bay, LONT06 for Little Salmon River, LONT07 for Stony Creek and LONT08 for Chaumont River (Figure 1.1 and Table 1.1).

1.3.1 Standard length and centroid size

The average standard length (SL) collected for round goby, *N. melanostomus* was similar among localities. The highest mean was collected from LONT02 and the lowest from LONT05 (Table 1.2). The SL recorded in the present study was relatively higher than the smallest reported mature individual caught in the Great Lakes, for example 45mm SL (MacInnis and Corkum 2000a; MacInnis and Corkum 2000b), suggesting the samples in the present study were of adult size. For pumpkinseed, *L. gibbosus* the lowest mean SL was from LONT04 and the highest from LONT02.

The relationship between SL and centroid size, CS was positively significant for round goby (CS= 20.24 + 7.88SL, $F_{1,274}$=220, adjusted R^2 = 0.44, p = <0.0001). The relationship between the SL and the CS was better than that recorded for pumpkinseed

(CS= 2.66 + 10.77SL, adjusted $R^2 = 0.80$, $F_{1,296}=1204$, $p = <0.0001$). The significant relationship between SL and CS suggested that the use of CS as an approximation of specimen size in shape analysis was appropriate.

Table 1.2. Sample size, standard length, SL in millimeters (mm) and centroid size, CS for the three study species.

Species	Site	Sample size (female, male)	Sex-pooled SL, Mean ± 1 SD mm
Round goby, *N. melanostomus*	LONT01	14, 9	$60 \pm 1.9^{a,b,c}$
	LONT02	11, 25	70 ± 2.9^{c}
	LONT03	21, 16	$65 \pm 1.7^{b,c}$
	LONT04	8, 32	68 ± 1.9^{c}
	LONT05	5, 10	$53 \pm 1.3^{a,b}$
	LONT06	14, 14	$61 \pm 2.0^{a,b,c}$
	LONT07	23, 25	$62 \pm 1.2^{b,c}$
	LONT08	22, 20	$63 \pm 1.3^{b,c}$
Pumpkinseed *L. gibbosus*	LONT01	27, 21	$89 \pm 2.3^{a,b,c,d}$
	LONT02	15, 23	107 ± 3.1^{c}
	LONT03	26, 23	106 ± 2.2^{c}
	LONT04	10, 12	$81 \pm 2.0^{a,b,d}$
	LONT05	18, 18	$105 \pm 1.8^{c,d}$
	LONT06	20, 21	$90 \pm 2.9^{a,b,d}$
	LONT07	9, 22	$98 \pm 2.6^{b,c,d}$
	LONT08	19, 21	$88 \pm 2.1^{b,c,d}$

Different superscript letters indicate means are in different group, tested using Tukey's method at alpha level = 0.05

1.3.2 Shape differences between localities

The interaction between locality, sex and centroid size was not significantly different for either species, indicating similar growth trajectories among localities and sexes for each species (Table 1.3). Therefore, multivariate analysis of covariate (MANCOVA) was carried out to test the response (shape) interaction between locality and sex. Locality and sex was significant for both species (Table 1.4). Despite that, for both species, the differences for location were twice the variance for sex for round goby (25% vs. 18%) and pumpkinseed (19% vs. 8.4%), indicating that the shape difference was more explained by locations than sex (Table 1.4). The interaction term for locality and sex was not significant for round goby, but was significant for pumpkinseed (round goby $F_{112,1542}$=1.096, p=0.240; pumpkinseed $F_{112,1542}$=1.210, p=0.035; Table 1.4). Because we were only interested in the variation of shape among localities rather than within localities, we pooled sex to increase sample size.

Table 1.3. Wilk's lambda test for parallel slope of locality and sex. The tests were performed on individual averaged Procrustes coordinates obtained after image superimposition.

Species	Effect	F	d.f. 1	d.f. 2	*p*-value	Variance %	Observed Power
Round goby,	Locality	1.17	112	1444.95	0.113	0.077	1.000
N. melanostomus	Sex	1.17	16	222	0.296	0.078	0.753
	Centroid size, CS	3.51	16	222	<0.001	0.202	1.000
	Locality x Sex	1.19	112	1444.95	0.098	0.078	1.000
	Locality x CS	0.79	112	1444.95	0.946	0.053	0.993
	Sex x CS	1.07	16	222	0.384	0.072	0.705
	Locality x Sex x CS	0.76	112	1444.95	0.973	0.051	0.989
Pumpkinseed,	Locality	1.02	112	1697.10	0.441	0.058	1.000
L. gibbosus	Sex	1.16	32	522	0.248	0.067	0.945
	Centroid size, CS	1.18	16	261	0.287	0.067	0.762
	Locality x Sex	0.96	176	2404.01	0.651	0.054	1.000
	Locality x CS	0.91	112	1697.10	0.733	0.053	0.998
	Sex x CS	1.18	32	522	0.227	0.068	0.949
	Locality x Sex x CS	0.94	176	2404.01	0.702	0.054	1.000

d.f. = degree of freedom

Table 1.4. Wilk's lambda test of multivariate analysis of variance.

Species	Effect	*F*	d.f. 1	d.f. 2	p-value	Variance %	Observed Power
Round goby	Locality	5.02	112	1541.93	<0.001	25.0	1.000
N. melanostomus	Sex	2.34	16	237	0.003	13.7	.986
	Locality x Sex	1.10	112	1541.93	0.240	6.8	1.000
	Centroid size	6.77	16	237	<0.001	31.4	1.000
Pumpkinseed	Locality	4.24	112	1826.40	<0.001	19.2	1.000
L. gibbosus	Sex	1.60	32	562	0.021	8.4	.993
	Locality x Sex	1.21	176	2586.50	0.035	6.4	1.000
	Centroid size	22.63	16	281.000	<0.001	56.3	1.000

d.f. = degree of freedom

1.3.3 Validation of locality 'membership'

We used the residuals of sex-pooled regression of Procrustes coordinates on centroid size for the subsequent analysis. Locality cross-validation revealed that 22-64% individuals of round goby were correctly classified, with the highest locality that was correctly classified was LONT02 (Table 1.5). For the pumpkinseed, the percentage of correctly-classified individuals ranged between 31 to 50% (Table 1.5). Similarly, the highest locality that was correctly classified was sample from LONT02. The misclassification rate for both species was relatively high, suggesting only subtle body shape variation among individuals across localities. We did not find any pattern in morphological shifts associated with the classification of individuals to each locality.

Table 1.5. Jack-knifed cross-validation of site-pair comparison analyzed in PAST software. Analysis was conducted using 'size-corrected' shape. Rows are the observed classification, while columns are cross-validated classification.

Species	Locality	LONT 01	LONT 02	LONT 03	LONT 04	LONT 05	LONT 06	LONT 07	LONT 08	Total	% correctly classified
Round goby	**LONT01**	5	0	1	5	2	3	4	3	23	22
	LONT02	1	23	5	2	2	1	2	0	36	64
	LONT03	4	7	9	4	1	2	8	2	37	24
	LONT04	4	5	3	18	2	4	3	1	40	45
	LONT05	3	2	0	1	6	1	1	1	15	38
	LONT06	7	0	1	5	1	11	3	0	28	39
	LONT07	2	3	8	5	1	3	19	7	48	40
	LONT08	4	2	5	4	3	2	3	19	42	45
	Total	30	42	32	44	18	27	43	33	269	
Pumpkinseed	**LONT01**	17	4	3	10	6	3	3	2	48	35
	LONT02	0	19	6	5	4	2	0	2	38	50
	LONT03	4	8	18	7	5	4	3	2	51	35
	LONT04	5	3	2	8	2	2	2	1	25	32
	LONT05	6	1	8	4	12	3	2	3	39	31
	LONT06	1	2	6	4	6	10	6	8	43	23
	LONT07	1	1	5	0	2	3	14	8	34	41
	LONT08	0	1	2	3	2	7	10	15	40	38
	Total	34	39	50	41	39	34	40	41	318	

1.3.4 Shape variation components

CVA was used to reduce the amount of shape variables from 20 variables (10 landmarks X *x*,*y* coordinates) into more simplified seven variables. The scores from the simplified shape variables (canonical variates scores, CV) were used to show the direction of separation of shape between localities.

For round goby, the direction in shape changes that was affected by size is presented in Figure 1.2. Therefore variation as depicted by CV1 (first component of CV) was confounded with size, even after the size removal through regression ($F_{1,267}$=10.08, p<0.002). CV1 component accounted for 38% of the total variance between localities, while CV2 accounted for 27% (Figure 1.2). Yet CV2 did not confound with size ($F_{1,267}$=1.766, p=0.185). The first two CV components accounted for two-thirds of total variation, leaving the rest of the CV's relatively unimportant source of variation. Most mean scores for CV2 center on the positive axis except for LONT03, LONT07 and LONT08. Morphological shifts of CV2 mainly occurred in caudal peduncle (compression of landmark 3,4,5 and 6; Figure 1.2), where individuals with positive CV scores would have shorter caudal peduncle than those with negative CV scores (Figures 1.2 and 1.4).

For pumpkinseed, morphological variation for the first canonical variate scores (component CV1 centers on the head region shift and caudal peduncle) accounted for 43% of total shape variation across samples (Figure 1.3). Interestingly the CV1 distribution data coincided with geographic position of the locations. Positive CV1

scores show longer distances between snout and the base of dorsal fin, shallower body depth and shorter caudal peduncle that was found for LONT06, LONT07 and LONT08 on the three farthest eastern locations of our samples (Figure 1.5). We also found no evidence for correlation between CV1 and size (CV1, $F_{1,316}$=1.22, p=0.2687). CV2 was associated with the expansion of body depth of pumpkinseed that accounted for 23% of the shape variation (Figure 1.3). CV2 was not correlated with size (CV2, $F_{1,316}$=3.238, p=0.0729). Our results revealed that CV2 is related to the habitat where the samples were collected. Samples collected from river mouth localities were on the negative side of CV2 axis as compared to bay samples that are on the positive size of CV2 axis. CV2 scores for LONT01, LONT03, LONT04 and LONT05 showed the samples had deeper body depth (relative distance between the base of first dorsal fin and pelvic and anal fins, Figure 1.5). Although CV3 account for 20% of variation, the component is strongly associated to specimen size ($F_{1,316}$=24.48, p<0.001) and therefore, like the rest of variation components was omitted from analysis due to relatively unimportant variation.

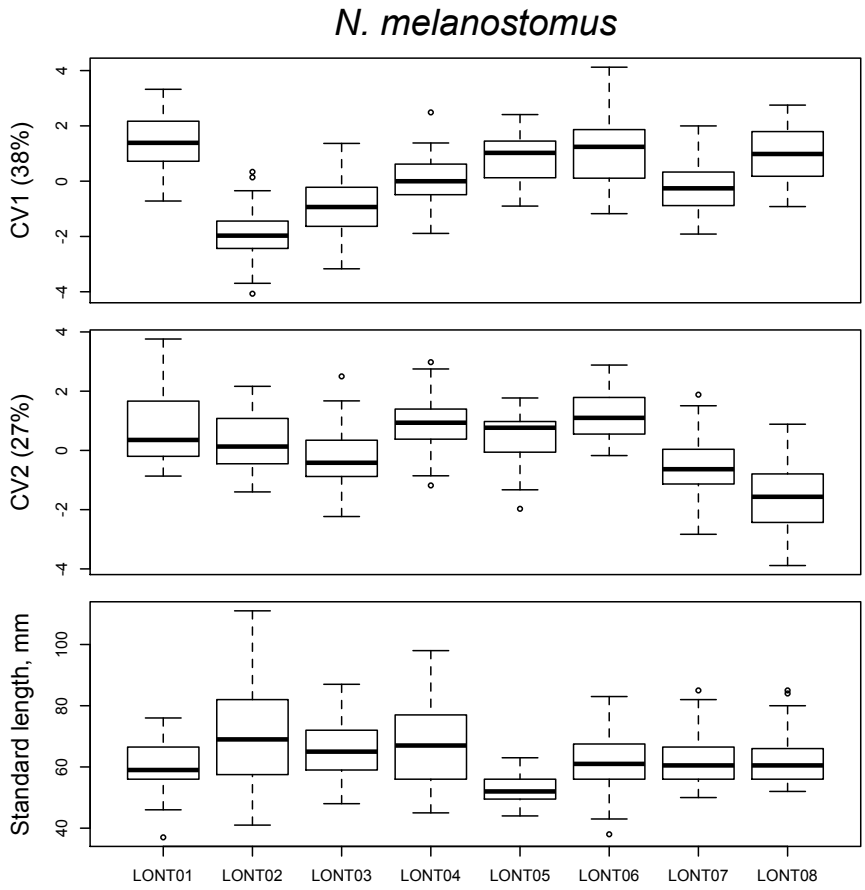


Figure 1.4. Boxplots of the first two canonical variate scores (CV) data distribution with associated percentage of variance in parentheses for round goby *N. melanostomus*. CV1 is associated with specimen size, while CV2 is associated with caudal peduncle variation (see Figure 1.2).

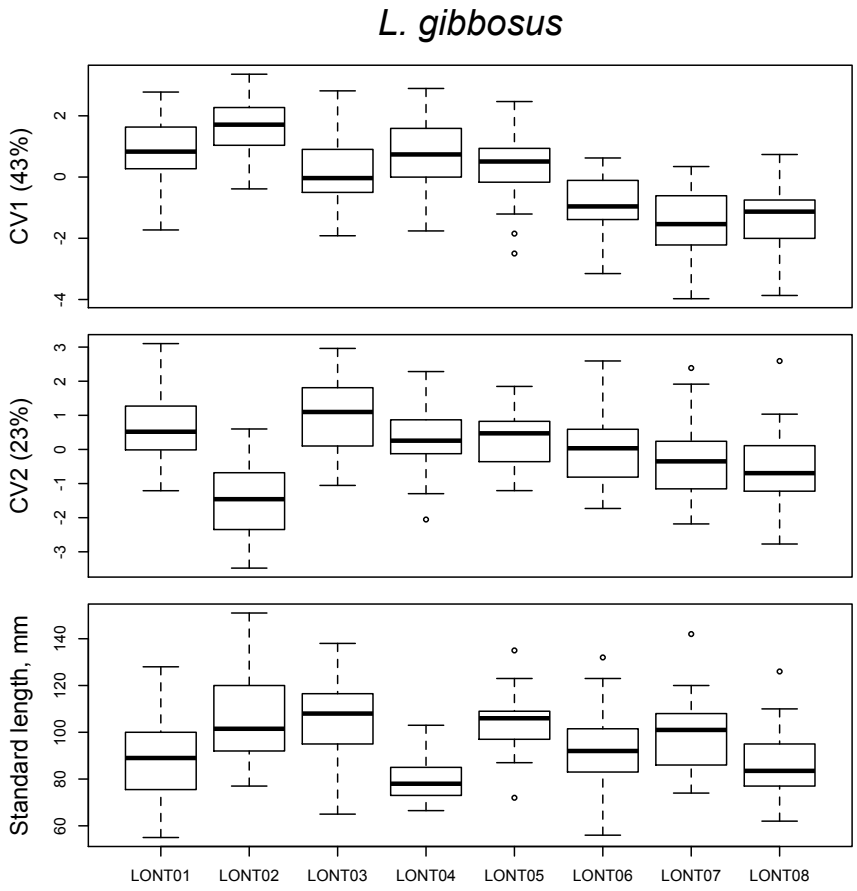


Figure 1.5. Boxplots of the first two canonical variate scores (CV) data distribution with associated percentage of variance in parentheses for pumpkinseed, *L. gibbosus*. CV1 is associated with specimen size, while CV2 is associated with caudal peduncle variation (see Figure 1.3).

1.3.5 Effect of geographic distance on morphological differences among species

Plotting the morphological differences between samples by means of Mahalanobis distance for each sample along an axis of geographic separation of localities revealed variable results for both species. We plotted the Mahalanobis distance of locality against associated geographic distance: (1) between all localities; (2) between bay and river mouth localities; (3) between bay localities; (4) between river mouth localities, for both round goby and pumpkinseed (Figure 1.6 and Figure 1.7). For round goby the geographic distance did not significantly predict the shape differences (see Figure 1.6 for equations, p values and R^2 values for the four localities vs. distance). Pumpkinseed showed significant relationship for all locality variation ($F_{1,26}$=11.94, p=0.0019) and between river mouth localities ($F_{1,26}$=7.749, p=0.0496) against the geographic distance Figure 1.7). The model essentially suggested as distance increased, the body shape variation was stronger between localities for pumpkinseed. We found that this relationship coincided with spatial structure in the morphological differences across the study area (Figure 1.5).

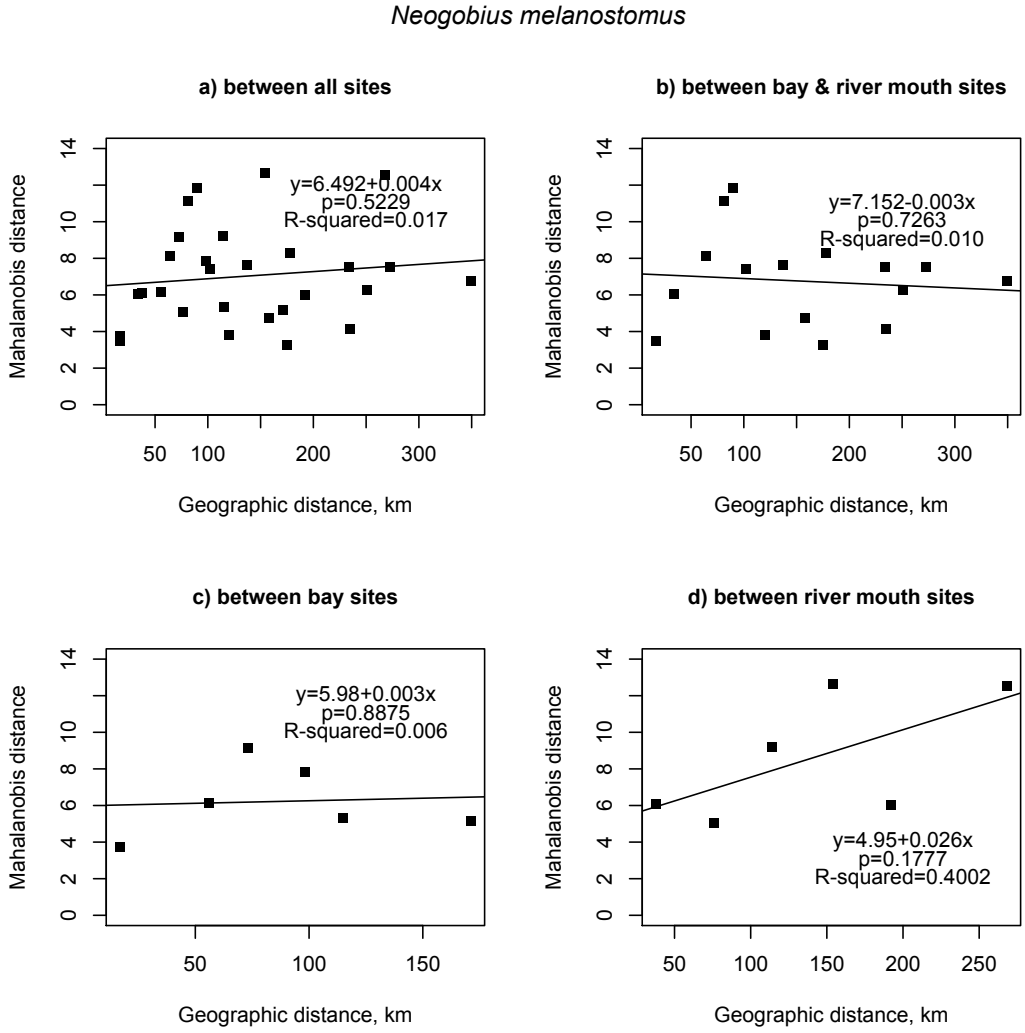


Figure 1.6. Sample pair-wise shape differences and geographic distances for round goby, *N. melanostomus*. None of the regression lines showed significant slopes.

Lepomis gibbosus

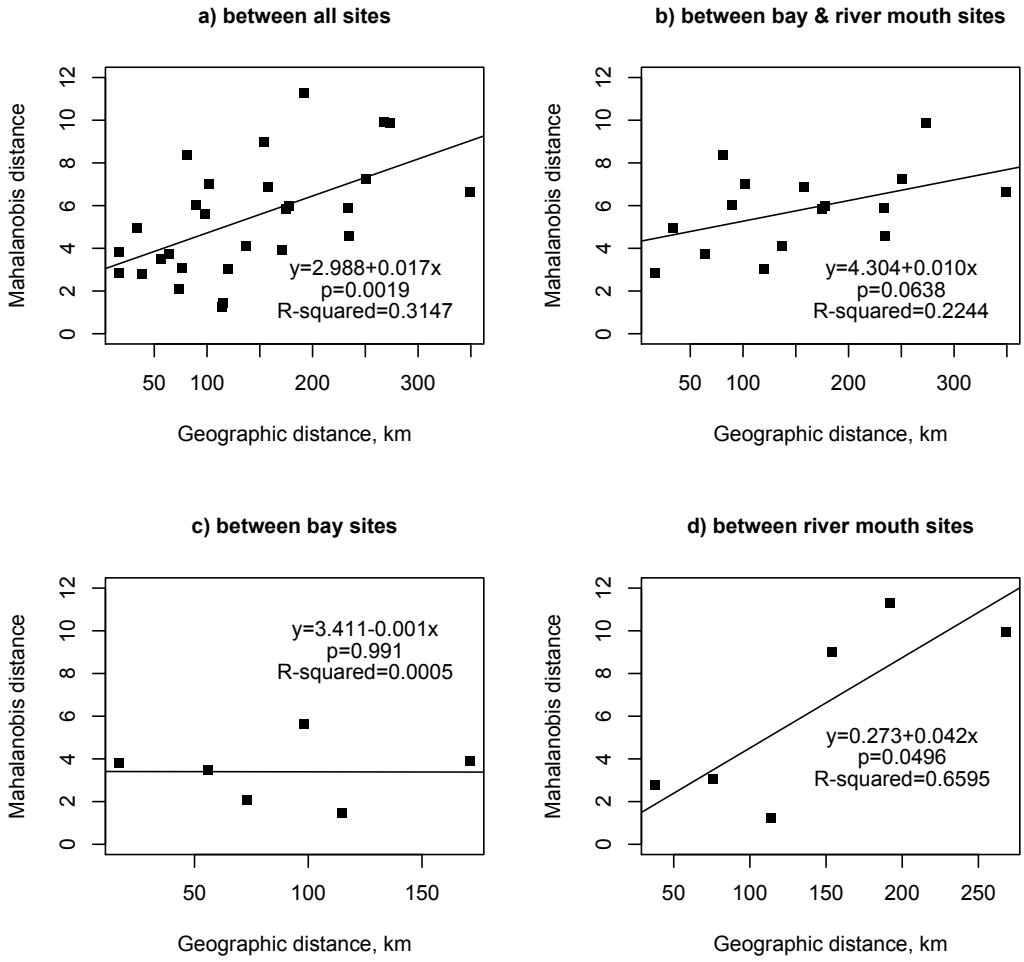


Figure 1.7. Sample pair-wise shape differences and geographic distances for pumpkinseed, *L. gibbosus*. Geographic distance significantly predicts morphological variation between (1) all sites and (2) river mouth sites.

1.4 Discussion

The mean morphological differences for round goby and pumpkinseed were too small to be visually identified but were identifiable with multivariate and ordination analyses. We observed major morphological differences in pumpkinseed for head length, caudal peduncle and body depth. The morphological variation coincided with geographic position of the localities along the shoreline of Lake Ontario, New York. In addition, the strength of the morphological differences was predicted by spatial distance among samples. For round goby, we found major morphological shifts such as elongation of posterior and anterior region of body shape were associated with size; however, spatial distance did not predict morphological differences.

In our present study, locations rather than habitats were chosen *a priori* to test whether distance limits population intermixing along the coastal line of Lake Ontario. Three of the eastern-most samples (sites LONT06, LONT07 and LONT08) were located adjacent to river outlets would receive highly oxygenated discharge from upstream (Randall et al. 1995). Fish inhabiting such fast-flowing water typically exhibit intra-specific variation related to locomotion such as shifts in body depth and placement of lateral fins as compared to their counterpart inhabiting much slower-moving current in lotic environment (Brinsmead and Fox 2002; Langerhans 2008). Nearby land development (sites LONT01 and LONT04) and land use for agriculture (site LONT02) can cause impervious surface cover in urban catchments, which alter hydrology and geomorphology of the streams through loading of nutrients, metals, pesticides and other contaminants into the waters that would result in decline in

richness of algal, invertebrates and fish communities (Paul and Meyer 2001).

1.4.1 Shape variation for round goby

We did not observe any significant variation of body depth for round goby from the bay and river mouth samples. Round goby has a fusiform shape that is adapted for swimming in high water velocity (Langerhans and Reznick 2010). Also round goby, have a pair of modified pelvic fins that functions like a suction disc that would be able to increase stability in fast-flowing water (Budney and Hall 2010). The torpedo-like body shape and the presence of pelvic fins in the form of a sucking disc enable round goby to adapt in a fast-flowing water condition despite the benthic behavior where one would presume individuals to experience a hydrodynamic deficiency (Tierney et al. 2011). The energy spent on the resistance might explain why round goby exhibit limited movement i.e. short home range (Marentette et al. 2012; Ray and Corkum 2001). However, lack of differences in the study area, as one would expect from limited dispersal, can be explained by short time period since introduction and various mechanisms of dispersal such as migration during larval stage and multiple introduction in the area (Brown and Stepien 2009; Hensler and Jude 2007) that can constrain divergence.

The unchanged body shape in round goby was also found in a previous study and which was attributed to development that represents a strongly precocial (i.e. specialized) life-history (L'avrinčíková, Kováč, and Katina, 2005). Early maturation and reproduction in the non-native population of round goby suggests a shift back

towards more altricial (i.e. less specialized, more generalized) life history (Corkum et al. 1998; L'avrincikova et al. 2005; MacInnis and Corkum 2000a; MacInnis and Corkum 2000b). A species with low genetic and constrained gene flow can show phenotypic differences in a few generations (fewer than 50) when either heritability or mixing are low (Hendry et al. 2001), a condition that an invasive species would likely experience. However, round goby in the invaded areas have high genetic diversity (Brown and Stepien 2009; Stepien et al. 2005) contrary to what a typical introduced species would experience in a new area and reproduce at faster rate regardless. The ability to adapt a wide range of habitats, a short elapsed time of population growth between arrival and establishment (Vélez-Espino et al. 2010) with altricial–precocial trajectories seems to be typical for an invasive species and may be one of key factors for their successful colonization of new environments (L'avrinčíková et al. 2005).

1.4.2 Shape variation for pumpkinseed

Differences in body form were evident from the CVA, despite some overlap among the samples from different sites. The degree of overlap on the canonical axes and the few fish at the extreme ends of the axes suggest that pumpkinseed do not fall into two discrete categories (i.e. bay and river-mouth morphs). Pumpkinseed samples collected from bay locations had deeper body, longer caudal peduncle and longer head than the river-mouth locations. Vøllestad et al. (2004) found that crucian carp *Carassius carassius* (L.) with deeper bodies had an increase in drag at swimming. Shallow body is a character of fish living in fast-flow stream habitat as opposed to deep body, a character for fish living in a slower-flow stream habitat. Another study related body

depth to schooling in the lakes and territoriality in the streams (Swain and Holtby 1989). Whereas another study indicated how diet affected body shape variation (e.g. body depth and tail length) in two species of *Geophagus* of family Cichlidae (Wimberger 1992).

Our result for pumpkinseed confirms an earlier study by Brinsmead and Fox (2002) on the effect of divergent selection on pumpkinseed which found deeper bodies in slow-moving water (LONT01, LONT03, LONT04 and LONT05) as compared with fast-flowing water in stream area (LONT2, LONT06, LONT07 and LONT8). We did not measure the body width, but Brinsmead and Fox (2002) associated fish with shallow body depth with slender body shape. As the burden of swimming increases in streams, due to the hydro-dynamic conditions, fishes selected for sustained swimming ability are generally more slender-bodied, rounder in cross-section and have a greater proportion of red muscle tissue whereas more sedentary lake fishes are generally more gibbose, more laterally compressed (or oblong in cross-section) and have a higher percentage of white muscle tissue (Brinsmead and Fox 2002; Lauder and Drucker 2004). Conversely fish with deep-bodied shape will add drag penalty that consequently decrease the swimming performance, although such minimization of resistance does not appear to be important in the type of low-speed maneuvering performed by fishes that forage in complex lake environments (Ralph et al. 2012).

1.4.3 The role of geographic distance in morphological variation

Geographic distance could limit population mixing and has been found previously to

be correlated with morphological divergence in two distantly related Neotropical fish between two habitats, river channel and lagoon sand banks (Langerhans et al. 2003). For pumpkinseed as the sampling locations were far away from each other, the degree of morphological variation in body depth became stronger. Studies from coastal systems suggest that near lakeshores can generate variation in the delivery of nutrients, phytoplantkons, macrobenthos and consequently fish assemblages. We expected no auto-correlation in terms of environmental variation in our study area because the shoreline is located along similar latitude that could otherwise vary gradually in temperature and ecosystem (Connolly et al. 2001). According to Langerhans et al., (2003), as long as the strength of divergent selection does not increase with geographic distance, the correlation between distance and divergence should explicitly test whether mixing constrains diversification.

For pumpkinseed, when locations from the western side alone (LONT01, LONT02, LONT03, LONT4 and LONT05) were analyzed, geographical distances were not significantly correlated with phenotypic differences. But when all samples were combined for a regression analysis, we observed a positive relationship between the degree of morphological differences and geographic distance for pumpkinseed, suggesting possible sub-divided population along the shoreline. The relationship indicated the role of spatial factor in maintaining divergence that could arise from limited population mixing in the area. However, as stated earlier, we were not able to determine whether the morphological sub-division is adaptive (genetic) or plastic (phenotype).

In the case of round goby, we hypothesize the lack of morphological differences in relation to spatial distance was due to: 1. they had little time to diverge since their arrival to Laurentian Great Lakes region in the late 1990's (Jude et al. 1992); 2. high population mixing during larval stage that could constrain differences or divergence (Hayden and Miner 2008; Hensler and Jude 2007); 3. the species has already developed traits that allow them to utilize different resources in a wide range of environmental conditions. We postulate that round goby has had higher population mixing (gene flow) than pumpkinseed, which requires further investigation.

Populations could diverge via alternative, genetically-based morphologies or through environmentally-induced phenotypes. The divergence among localities could be due to selective pressures from several factors such as predation and resources (Bartels et al. 2012). Such factors could in turn affect gene flow or population mixing in which gene flow should only constrain the magnitude of genetic differences, not plastic diversification (Mittelbach et al. 1999). Especially for pumpkinseed, the geographic arrangement (east-west) of the first CV component leads us to hypothesize that there is genetic differentiation that caused the morphological differences, while the second CV component that coincided with habitat types would lead us to hypothesize that there are also environmentally-induced effects in the morphological variation.

1.5 Conclusion

Consideration of geographic distance revealed the two species responded to this factor

in different ways, where the degree of morphological variation increased with distance for pumpkinseed but did not affect round goby. The extent of such morphological differences could stem from restricted dispersal in pumpkinseed, but not in round goby. Identification of the differential response in morphology suggests that localized threat on limited-dispersal species like pumpkinseed by invasive generalist species like round goby could further affect selection, which needs further investigation. For round goby, non-significant variation body shape among locations in our study is perhaps a character of a successful invader (L'avrincikova et al. 2005). It is also possible that selective pressures are resulting in different phenotypes on round goby. Coincidently, the pumpkinseed is an invasive species in Europe (Bhagat et al. 2011). Studies between native and invasive pumpkinseed populations inhabiting fluvial and lacustrine waterbodies showed that pumpkinseed have acquired adaptive external morphologies in the invaded range related to locomotion such as median-fin size and placement and also body width through successive generations (Yavno et al. 2012). In the invaded range where evolutionarily related species is not present, comparison with native species that is distantly related could provide an alternative way in understanding ecological causes of phenotypic patterns if both species converge in similar ecology-phenotype associations. Apart from morphological differences, we also demonstrated that spatial distance might serve as an important constraint on adaptive diversification. We suggest experimental studies to be carried out to examine the effect of divergent selection so as to further quantify the potential evolutionary effects of native and invasive species in the Laurentian Great Lakes region.

CHAPTER 2

OCCUPANCY OF ROUND GOBY, *Neogobius melanostomus* AND NATIVE BENTHIC SPECIES IN NEW YORK STATE TRIBUTARIES

Abstract

Knowledge of the spread of invasive species and factors that could potentially stave off their further invasion is important to consider when developing management plans. I used co-occupancy modeling to estimate the occurrence of benthic fish that namely round goby *Neogobius melanostomus* and darters in streams that drain into Lake Ontario, New York. I found that models that conditioned the occupancy of darters on the presence of round goby received stronger statistical support compared to models that did not. The results further showed that models that assumed perfect detection ($p = 1$) when round goby are present can be underestimated as much as 5 to 60% when compared with models that assumed imperfect detection ($p < 1$) in terms of the probability of occupancy of darters. While the occupancy of streams by round goby decreased with link magnitude (a proxy for stream size), the probability of occupancy by round goby can also be underestimated by 2 to 15% when imperfect detection is not accounted for. The study revealed that estimates of benthic fish occupancy are likely influenced by imperfect detection. The results are useful as a framework for improving investigations into questions of persistence and extirpation of native species when invasive species have already become established.

2.1 Introduction

Successful establishment of non-native species in a new ecosystem has often been reported to cause adverse impacts on ecosystems and economies. Round goby *Neogobius melanostomus* is an invasive species within the Great Lakes basin of North America, accidentally introduced through ballast water transport from Black, Caspian and Azov seas (Marsden et al. 1996). This species gained much attention from public and scientific communities in the recent years due to its rapid spread and level of invasiveness. Since its first sighting in 1990 in St. Clair River, Michigan (Jude et al. 1992), the round goby population has successfully expanded and has become established across the other Great Lakes and into the tributary streams (Clapp et al. 2001; Kornis and Vander Zanden 2010; Pennuto et al. 2010; Phillips et al. 2003). As more and more areas are being invaded, knowledge about the spread of round goby and its possible effects on other species is important to consider for future conservation plans.

In tributary streams, small native benthic fishes like darters of genera Etheostoma and Percina are the most likely to be affected by the round goby invasion through competition (Balshine et al. 2005; Bergstrom and Mensinger 2009; Dubs and Corkum 1996; French III and Jude 2001) owing to their benthic morphology that lack swim-bladders. Round goby invasion has been shown to reduce the abundance of these small benthic fishes (Phillips et al. 2003) and in one study, eventually led to local extinction of native benthic species (Janssen and Jude 2001). Low abundance and elusive behavior may have caused these species to not always be detected in the locations they

occupy (Beneteau et al. 2011; Groce et al. 2012; Sterling et al. 2012). Also, the lack of detection could be due to allopatric distribution that is prevalent among darters (Page and Schemske 1978; van Snik Gray et al. 2005; Snik Gray and Stauffer 2001) as well as distribution barriers and historical origins of species (Byholm et al. 2012). Additionally, round gobies are more aggressive, exhibit a wide range of habitat tolerance, are highly abundant once present (Bergstrom et al. 2008; Coulter et al. 2012; Lynch and Mensinger 2012; Pennuto et al. 2010) and therefore are possibly easier to detect when present. Thus, using models to estimate the distribution of both round goby and darters without accounting for disparities in species detection could limit our interpretation of the magnitude and direction of species distributional interaction. Consequently, this may lead to inaccurate assessment in the development of conservation management plans.

In the present study, we attempted to estimate the co-occupancy of round goby and darters by accounting imperfect or false detection that could arise from the factors discussed in previous paragraph. I hypothesize: (1) that failure to account for imperfect detection probabilities will bias the estimates of benthic fish occupancy, (2) that the aquatic landscape (stream size and distance from shore) can be used to predict occupancy of benthic species. In examining the predictions, we used a likelihood-based occupancy model (Mackenzie et al. 2004) that accounts for false absences, which was later extended into a two-species occupancy model that characterizes species interactions while accounting for habitat covariates (Richmond et al. 2010).

2.2 Materials and Methods

2.2.1 Study area

The study area in New York state spans the northwestern watershed region that discharges into Lake Ontario (Hydrography Unit Code, HUC-12, 04130001 0902, 0901, 0802, 0801, 0704, 0603, 0602, 0601 and 0504). The maximum size of stream order considered was 5, with stream link magnitudes ranging from 1 to 51.

Site Selection

Site selection was limited to an area within a U.S. Geological Survey (USGS) 10-digit Hydrologic Unit Code (HUC) of the western side of New York's Lake Ontario watershed that consisted nine 12-digit HUC, allowing for greater sampling effort per area given other logistical constraints. Impoundments layers were plotted from two sources: the national hydrography datasets (hereafter, NHD) and the national inventory of dams. All sites upstream of a dam or waterfall without a lock to mediate transport were omitted from further selection, because we assumed that these barriers would impede round goby's natural dispersal into upper-stream areas (Kornis and Vander Zanden 2010). All easily accessible stream sites (e.g., road crossings, trails, easy boat access) were plotted based on a 1:100,000 NHD. Within each of the nine 12-digit HUC watersheds, four accessible sites were randomly selected using a stratified random selection design based on stream size classification (i.e. link magnitude; Shreve 1966). Stream link magnitude is defined as the number of un-branched source streams upstream from a given segment in the drainage network.

A total of 194 sites with stream link magnitudes ranging from 1 to 51 were obtained from the study area. Sampling was conducted by a field crew of four. After reconnaissance site surveys and pilot sampling, we specifically excluded sites with stream link magnitudes under 4 because our visits during the summer showed that sites below this size became dry. Within each 12-digit HUC watersheds, we stratified site selection by randomly selecting two sites with stream link magnitude 4-8, one 9-15-link streams and one above 16-link streams. This criterion was designed to ensure we sampled various stream sizes and habitat types so that the chosen sites were a relatively unbiased representation of the watershed. In total, 39 sites were selected for the study (Figure 2.1).

2.2.2 Sampling of round goby and darter occurrence

Sampling was conducted in June and July of 2011. Because the heterogeneity in geographic condition of streams i.e. shallow vs. deep, we sampled using different gear types depending on depth of water. At shallow-water sites (wadable at about 1.2m-deep) we used a beach seine net and stationary electrofishing, whereas at deep-water sites we used boat-electrofishing and angling at deep-water sites (Appendix 2.1 for sampling procedure details).

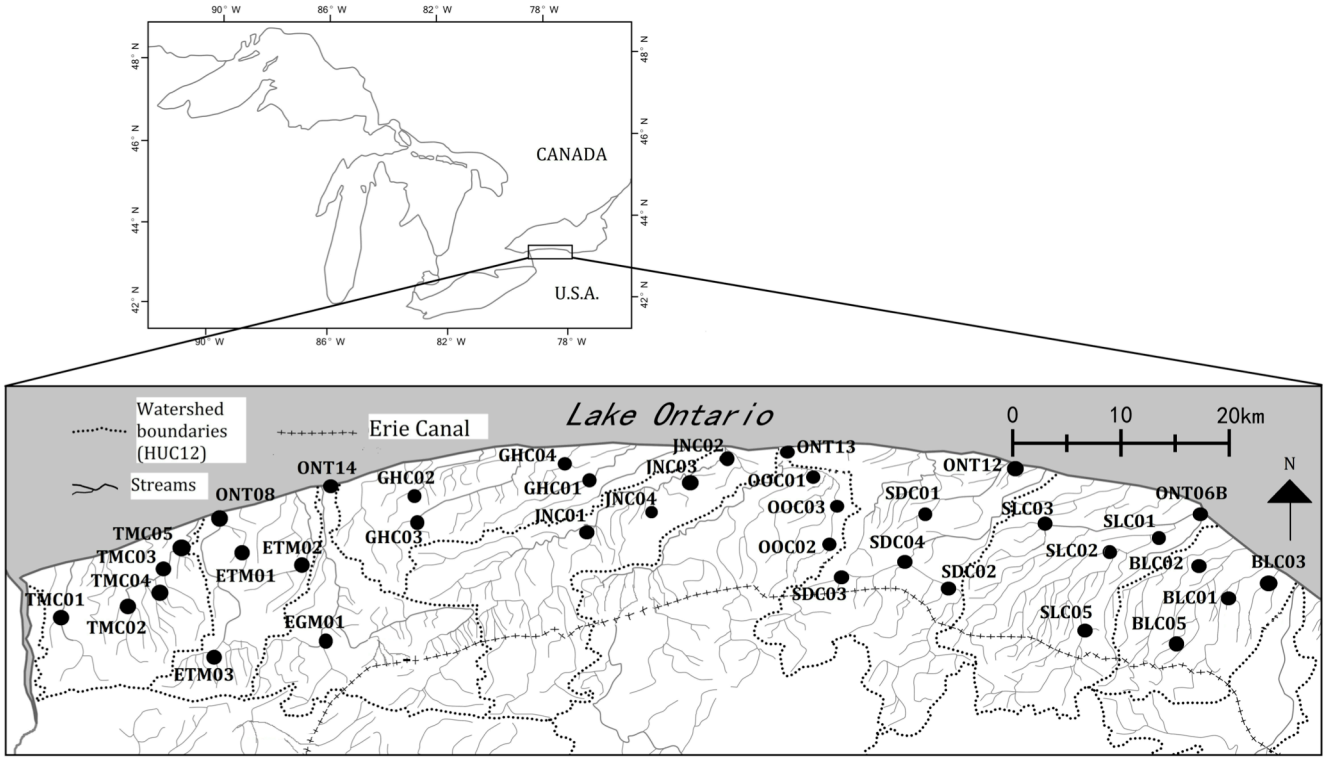


Figure 2.1. Map of study area in 12-digit subwatershed region from National Hydrographic Dataset (NHD HUC12).

In shallow water sites, the time gap between the use of seine gear and stationary electrofishing was about one hour in order to ensure the site returned to its normal state prior to sampling activities. We conducted sampling without removal for the first sample and collected round goby after second sample for use for additional study. In deep-water sites, we employed both gears (angling and boat electrofishing) for about an hour simultaneously. We saved any round goby collected during both kinds of sampling. During all sampling, field crews identified the presence of round goby as well as native species. Detected fish were enumerated and identified to species level for darters and to lowest taxa for other group of fish in the field and released.

Covariates for site occupancy and sampling detection

At shallow water sites, we recorded dominant substrate type (cobble, boulder, silt, sand) for each sampling quadrat and took the percentage of substrate type over all substrate types determined during seining, as a measure of substrate type for each site. We recorded the depth (to nearest 0.1cm), using a depth pole, by measuring averaging depth of three relatively equally spaced points at a site. We measured the wetted width (i.e., bank-to-bank distances at water level) with a measuring tape at approximately widest and narrowest (to nearest 0.1m) area. Additionally, we enumerated all of the species to the lowest taxa encountered in each quadrat. I used Google Earth to measure proximate distance as the fish swims, at the middle of the stream width between each site and their respective stream mouth to Lake Ontario.

Model assumption

We assumed sites to be closed to changes in occupancy state because round goby and darter are sedentary species and their local movement/migration are small enough to not affect their occupancy status given the small time window of the sampling period.

2.2.3 Estimating the round goby and darter co-occupancy

I used a conditional parameterization of the two-species occupancy model to test the hypothesis that the state of darter occupancy was affected by the presence or absence of round goby as influenced by covariates. Following parameterization for the model outlined in Richmond et al. (2010), for each site there are four possibilities to estimate co-occupancy states for round goby and darter (Appendix 2.2 for model structure): (1) both round goby and darter species were present (ψ^{DG}); (2) only round goby was present ($1-\psi^{DG}$); (3) only darter species were present (ψ^{Dg}); or (4) neither species is present ($1-\psi^{Dg}$). The parameters and associated descriptions are listed in Table 2.1. Site occupancy was estimated from detection histories that consisted of sequences of detections (1) and non-detections (0) for each sampling occasion at 39 sites for both round goby and darters (Table 2.2; Appendix 2.3 for detection model description).

Two-species occupancy models using these parameters in this study lead to 20,419 models (occupancy, 3^9 + detection, 3^6) representing the different combinations of covariates needed to test my hypothesis (conditional and unconditional parameterization for occupancy, detection and covariates). Because of having so many models to run, I restricted model selection by running biologically reasonable

candidate models (see Table 2.3 for full list of occupancy covariates, and covariate X species interactions used in the two-species occupancy models) while allowing only link magnitude to be included as a covariate in the detection models after testing for the best detection covariate (Appendix 2.4 for methods determining the best detection covariate).

Table 2.1. Descriptions of the parameters used in the round goby and darter occupancy model following Richmond et al. (2010).

Parameter	Description
ψ^{G}	Probability of occupancy for round goby
ψ^{DG}	Probability of occupancy for at least one darter species, given round goby are present
ψ^{Dg}	Probability of occupancy for at least one darter species, given round goby are absent
p^{G}	Probability of detection for round goby, given all darter are absent
p^{D}	Probability of detection for at least one darter species, given round goby is absent
r^{G}	Probability of detection for round goby, given both round goby and a darter species are present
r^{DG}	Probability of detection for at least one darter species, given both round goby and at least one darter species are present and round goby is detected
r^{Dg}	Probability of detection for at least one darter species, given both round goby and at least one darter species are present and round goby is not detected

Table 2.2. Encounter history for each sample point, from which detection probability of round goby and darters calculated. Subscripts represent sampling.

Species		Frequency of encounter history
Round goby	**Darters**	
$_1$1 $_2$1	$_1$1 $_2$1	8
$_1$1 $_2$1	$_1$1 $_2$0	6
$_1$1 $_2$1	$_1$0 $_2$1	0
$_1$1 $_2$1	$_1$0 $_2$0	5
$_1$1 $_2$0	$_1$1 $_2$1	1
$_1$1 $_2$0	$_1$1 $_2$0	0
$_1$1 $_2$0	$_1$0 $_2$1	0
$_1$1 $_2$0	$_1$0 $_2$0	0
$_1$0 $_2$1	$_1$1 $_2$1	1
$_1$0 $_2$1	$_1$1 $_2$0	2
$_1$0 $_2$1	$_1$0 $_2$1	0
$_1$0 $_2$1	$_1$0 $_2$0	0
$_1$0 $_2$0	$_1$1 $_2$1	7
$_1$0 $_2$0	$_1$1 $_2$0	0
$_1$0 $_2$0	$_1$0 $_2$1	0
$_1$0 $_2$0	$_1$0 $_2$0	9
	Total	39 sites

i = sampling history

Table 2.3. Relative support for two-species occupancy models for darters and round goby with occupancy covariates and covariate X species interaction terms. w_i = Akaike information criterion (AIC) weight of the evidence.

Site covariate	***N***	w_i
LMAG, LMAG X Species (U), DIST, DIST X Species (C)	2	0.252
LMAG, LMAG X Species (U), DIST, DIST X Species (U)	4	0.225
LMAG, LMAG X Species (C), DIST, DIST X Species (C)	2	0.216
LMAG, LMAG X Species (C), DIST, DIST X Species (U)	2	0.214
LMAG, LMAG X Species (U), DIST	4	0.070
LMAG, LMAG X Species (C), DIST	2	0.007
LMAG, LMAG X Species (C), DIST, DIST X Species (C), SUBS, SUBS X Species (C)	2	0.002
LMAG, DIST, DIST X Species (U)	4	0.001
LMAG, LMAG X Species (U), DIST, DIST X Species (U), SUBS, SUBS X Species (U)	2	0.000
LMAG, DIST	4	0.000
LMAG, DIST, DIST X Species (C)	2	0.000
LMAG, SUBS	4	0.000
LMAG, LMAG X Species (U), SUBS	4	0.000
LMAG, LMAG X Species (C), SUBS	2	0.000
DIST	4	0.000
DIST, DIST X Species (U)	4	0.000
DIST, DIST X Species (C)	2	0.000
LMAG	4	0.000
LMAG, LMAG X Species (C)	2	0.000
LMAG, LMAG X Species (U)	4	0.000
DIST, SUBS	4	0.000
DIST, DIST X Species (C), SUBS	2	0.000
DIST, DIST X Species (U), SUBS	4	0.000
SUBS	4	0.000
SUBS, SUBS X Species (C)	2	0.000
SUBS, SUBS X Species (U)	4	0.000
(.)	4	0.000
DIST, SUBS, SUBS X Species (U)	2	0.000
LMAG, SUBS, SUBS X Species (U)	2	0.000
DIST, DIST X Species (U), SUBS, SUBS X Species (U)	2	0.000
LMAG, LMAG X Species (U), SUBS, SUBS X Species (U)	2	0.000

2.2.4 Two-species occupancy model with imperfect detection

I used the best detection covariates for each darter and round goby respectively in the two-species occupancy and detection models to simultaneously test:

1. whether the detection probability of darter was conditional on the presence or detection of round goby, where detection models for two-species occupancy were parameterized as p^{D} and $r^{DG} = r^{Dg}$ vs. $p^{D} = r^{DG} = r^{Dg}$;
2. whether darter occupancy was conditional on the presence of round goby, where ψ^{DG} and ψ^{Dg} were estimated separately i.e., the presence of darter was conditional on the presence of round goby compared to when $\psi^{DG} = \psi^{Dg}$ i.e., the presence of darter was unconditional;
3. whether occupancy for darter and round goby was influenced by covariates.

I predicted that the detection of round goby would have no effect on the detection probability of darter when both were present. This because the sampling gear used was not species-specific such that they will not capture only a specific species. Indeed we found darters in the same quadrat with round goby during our sampling. However, if round goby act as a dominant competitor, its presence was expected to reduce darter detection probability and/or occupancy.

I examined the effects of stream link magnitude (LMAG), distance from Lake Ontario (DIST) and percentage of hard substrates (SUBS) on occupancy for both species. I also fit models with LMAG times Species and DIST times Species interactions that tested whether darter occupancy changes as a function of the covariates conditional on

the presence of round goby. Like the occupancy models, this essentially means both LMAG and DIST were estimated in conditional/unconditional scenarios as well (see Table 2.3 for list of covariates tested). For instance for LMAG, the covariate was estimated as follows:

1. unconditional of both species being present or absent,

   $[(\psi^{G}, \psi^{DG}$ and $\psi^{Dg})$x(LMAG)];

2. unconditional on round gobies being present,

   $[((\psi^{G})$x(LMAG))+$((\psi^{DG} = \psi^{Dg})$x(LMAG))];

3. conditional on round gobies being present,

   $[((\psi^{G})$x(LMAG))+$((\psi^{DG})$x(LMAG))+$((\psi^{Dg})$x(LMAG))]

I also fit additive models that included combinations of covariates and species interactions.

2.2.5. Two-species occupancy model with perfect detection

I also ran the best conditional and unconditional models of two-species occupancy in the section 2.2.4 but with the detection models were held constant (.), which essentially combined the observation into one-observation, instead of two-observation history. The models were run to test the effect for not accounting for false detection in estimating species occurrence i.e. two-species occupancy models with perfect detection.

2.2.6. Model selection

Altogether I ran a candidate set of 92 two-species occupancy models (Appendix 2.5 for full list of models ran). For model selection, I used the difference between the Akaike information criterion (AIC) value for the model of interest and that for the best-supported model to assess the relative model fit (ΔAIC; Burnham and Anderson, 2004). The models were ranked based on ΔAIC, with ΔAIC = 0 being the best-supported model. I also present the corresponding weight of the evidence (w_i) of the ΔAIC, which is the ratio of ΔAIC of model of interest over the sum of all ΔAIC. The w_i can be interpreted as the probability that the model is the best approximating model given the data and the models examined (Akaike 1978).

2.2.7. Species interaction

I determined species interactions related to occupancy and detection probability as a function of stream link magnitude and distance from Lake Ontario using the best occupancy estimates for each species. The following equations are the estimates of the species interaction factor:

$$\varphi = \frac{\psi^{DG}}{(\psi^{G}\psi^{D})}$$

where, $\psi^{D} = (\psi^{G}\psi^{DG} + (1 - \psi^{G})\psi^{DG})$

The mean ± one standard error (s.e) are presented. We ran the single-season single-species and two-species occupancy models in Program PRESENCE, version 5.1 (Hines 2006).

2.3. Results

3.3.1 Round goby and darter occurrence

A total of 39 sites were sampled. Of these, 33 sites sampled with seine nets, 28 with stationary electrofishing, six with boat electrofishing and 11 by angling (Table 2.4; Appendix 2.6. for sampling information details). Distance between sites and the Lake Ontario shore varied from 0.6 to 22.4 km with a mean of 7.75 ± 7.03 km (mean ± s.d), stream link magnitude varied from 4 to 51 with a mean of 17 ± 15 km (mean ± s.d) and the mean percent of hard substrate at quantified sites varied from 3 to 90%. A total of 51 species from 13 families were detected during sampling (Table 2.4 and Appendix 2.7). Besides round goby, overall we found seven darter species: greenside darter *Etheostoma blennoides,* rainbow darter *E. caeruleum*, Iowa darter *E. exile*, fantail darter *E. flabellare*, tessalated darter *E. olmstedi*, logperch *Percina caprodes* and blackside darter *P. maculata* in the sampling area (Table 2.4). Because there was a sign of localized distribution of darter species, I combined the darters into a single darter category in the occupancy models to account for non-overlapped distribution.

Table 2.4. Round goby and darters distribution in the study area (D, species detected during at least one survey; ND, not detected at all surveys). DIST = distance from Lake Ontario, km; LMAG = stream link magnitude; SUBS = hard substrate type, %. †See Appendix 2.1 for details on associated streams.

Site†	Covariates			Species							
				RGB	Darter						
	DIST	LMAG	SUBS		GRN	RBW	IOW	FNT	TES	LOG	BLK
TMC01	2.2	7	90	ND	ND	ND	ND	ND	ND	ND	ND
TMC02	9.1	5	83	ND	ND	ND	ND	ND	ND	ND	ND
TMC03	2	16	92	ND	ND	ND	ND	ND	ND	ND	ND
TMC04	5.4	11	93	ND	ND	ND	ND	ND	ND	ND	ND
TMC05	0.4	16	90	D	ND	ND	ND	ND	ND	D	ND
ETM01	1.4	19	81	D	ND	ND	ND	ND	D	ND	ND
ETM02	6.9	4	25	ND	ND	ND	ND	ND	ND	ND	ND
ETM03	14.7	6	48	ND	ND	ND	D	ND	ND	ND	ND
ONT08	0.3	20	NA	D	ND	ND	ND	ND	ND	ND	ND
EGM01	14	9	90	ND	ND	ND	ND	ND	ND	ND	D
ONT14	0.5	23	NA	D	ND	ND	ND	ND	ND	ND	ND
GHC01	3.6	5	84	ND	ND	ND	ND	ND	ND	ND	ND
GHC02	3.4	4	90	D	ND	ND	ND	ND	ND	ND	ND
GHC03	4.5	4	90	ND	ND	ND	ND	ND	ND	ND	ND
GHC04	1.2	7	87	ND	ND	ND	ND	ND	ND	ND	ND
JNC01	9.4	37	90	D	D	ND	ND	ND	ND	ND	ND
JNC02	2.3	44	90	D	ND	ND	ND	ND	ND	D	D
JNC03	3.8	44	3	D	D	ND	ND	ND	D	ND	D
JNC04	7.6	43	90	D	D	ND	ND	ND	D	ND	D
OOC01	4.5	10	90	ND	ND	D	ND	D	ND	D	ND
OOC02	8.2	7	94	ND	ND	D	ND	D	D	ND	ND
OOC03	6.5	10	90	ND	ND	D	ND	D	D	ND	ND
ONT13	2.2	13	NA	D	ND	ND	ND	ND	ND	D	ND
SDC01	16.2	46	90	D	ND	D	ND	D	ND	ND	ND
SDC02	22.4	13	77	D	ND	ND	ND	ND	ND	ND	ND
SDC03	18	4	90	D	ND	D	ND	ND	D	ND	ND
SDC04	27	6	54	D	ND	D	ND	ND	D	ND	ND
ONT12	0.3	51	NA	D	ND	ND	ND	ND	ND	ND	ND
SLC01	5.2	32	89	D	ND	D	ND	D	D	D	ND
SLC02	8.7	20	90	D	D	D	ND	D	D	D	ND
SLC03	12.1	4	63	ND	D	D	ND	D	D	D	ND
SLC04	21.4	5	90	ND	ND	D	ND	D	D	ND	ND
SLC05	20	12	90	D	ND	ND	ND	D	ND	D	ND
ONT06B	1.2	45	NA	D	ND	ND	ND	ND	ND	ND	ND
BLC01	6.2	4	36	ND	ND	ND	ND	D	D	ND	ND
BLC02	4.8	9	79	D	ND	ND	ND	ND	ND	D	ND
BLC03	2	7	50	D	ND	ND	ND	ND	D	ND	ND
BLC04	0.6	37	56	D	ND	ND	ND	ND	D	D	ND
BLC05	12.2	4	90	D	ND	ND	ND	ND	ND	D	ND

ψ^{G} = Probability of occupancy for round goby
ψ^{DG} = Probability of occupancy for at least a darter species, given round goby are present
ψ^{Dg} = Probability of occupancy for at least a darter species, given round goby are absent
RGB = round goby, GRN = greenside darter, RBW = rainbow darter, IOW = Iowa darter, FNT = fantail darter, TES = tesssalated darter, LOG = logperch, BLK = blackside darter

2.3.3 Round goby and darter co-detection

Darter detection probabilities appeared to be conditional on round goby detection when both species were present. Overall, models that examined the relationship between darter detection probability when round goby was present, p^{G},p^{D},r^{D}, received better support as compared to models that assumed no relationship for both species group detection, p^{G},p^{D} (Table 2.5). As such, models that account for the conditional detection of darter on the presence of round goby had a total cumulative Akaike weight of 0.915 (Table 2.5). In total, conditional models that did not account for the effect of round goby presence on darter detection showed about 10 times lower accumulated support as compared to conditional models when the effect was accounted for (Akaike weight of 0.084 vs. 0.903; Table 2.6).

The results revealed much stronger support for darter detection probabilities when round goby was present (detection probabilities estimation of p^{D} and $r^{DG} = r^{Dg}$ received better support than that of $p^{D} = r^{DG} = r^{Dg}$; Model C01 vs C06; Table 2.5). The conditional detection probabilities for darter that was conditional on the presence of round goby, p^{D} and $r^{DG} = r^{Dg}$ (logit estimates, $\hat{e}$ = 27.9 ± >10, $\hat{e}$ ± 1s.e) yielded lower average detection probabilities of 0.72 that ranged from 0.70 to 0.77 ($\hat{e}$ = 0.96 ± 0.38, $\hat{e}$ ± s.e) than the unconditional detection probabilities ($p^{D} = r^{DG} = r^{Dg}$), that had an average of 0.79, ranging from 0.76 to 0.81 (1.38 ± 0.37, $\hat{e}$ ± s.e). The effect of link magnitude did not significantly affect detection probabilities for either detection model parameterizations (conditional and unconditional detection probability was 0.11 ± 0.31 and -0.10 ± 0.30 respectively, $\hat{e}$ ± s.e, $\alpha = 0.05$).

Table 2.5. The 20 top-supported two-species occupancy models examining interactions between darters (*Etheostoma* spp. and *Percina* spp.) and round goby (*Neogobius melanostomus*) out of a total set of 180 models. Models in italic are the best support models for conditional and unconditional occupancy respectively. Models in bold are the reciprocal of that best support models when detection probability assumed = 1. Link magnitude was used as a covariate for all detection models. ΔAIC = the difference between the Akaike information criterion (AIC) of model of interest and the best-supported model. w_i = AIC weight of the evidence.

Model code ()*	Occupancy model	Occupancy covariates	Detection model	ΔAIC	w_i
C01 (12)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	*LMAG, LMAG X Species (U), DIST, DIST X Species (C)*	p^{G}, p^{D}, r^{D}	*0.00*	*0.231*
C02 (11)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (U), DIST, DIST X Species (U)	p^{G}, p^{D}, r^{D}	0.24	0.205
C03 (13)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST, DIST X Species (C)	p^{G}, p^{D}, r^{D}	0.30	0.199
C04 (12)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST, DIST X Species (U)	p^{G}, p^{D}, r^{D}	0.32	0.197
C05 (10)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (U), DIST	p^{G}, p^{D}, r^{D}	2.55	0.065
C06 (11)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (U), DIST, DIST X Species (C)	p^{G}, p^{D}	4.81	0.021
C07 (10)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (U), DIST, DIST X Species (U)	p^{G}, p^{D}	4.93	0.020
C08 (11)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST, DIST X Species (U)	p^{G}, p^{D}	5.20	0.017
C09 (12)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST, DIST X Species (C)	p^{G}, p^{D}	5.21	0.017
C10 (9)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (U), DIST	p^{G}, p^{D}	7.64	0.005
C11 (12)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST	p^{G}, p^{D}, r^{D}	7.84	0.005
U01 (12)	$\psi^{G}\ \psi^{D}$	*LMAG, LMAG X Species (U), DIST, DIST X Species (U), SUBS, SUBS X Species (U)*	p^{G}, p^{D}, r^{D}	*8.24*	*0.004*
U02 (9)	$\psi^{G}\ \psi^{D}$	LMAG, LMAG X Species (U), DIST	p^{G}, p^{D}, r^{D}	8.32	0.004
C12 (10)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST	p^{G}, p^{D}	9.18	0.002
U03 (10)	$\psi^{G}\ \psi^{D}$	LMAG, LMAG X Species (U), DIST, DIST X Species (U)	p^{G}, p^{D}, r^{D}	9.33	0.002
U04 (9)	$\psi^{G}\ \psi^{D}$	LMAG, DIST, DIST X Species (U)	p^{G}, p^{D}, r^{D}	9.77	0.002
C13 (13)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, LMAG X Species (C), DIST,DIST X Species (C), SUBS, SUBS X Species (C)	p^{G}, p^{D}, r^{D}	9.89	0.002
C14 (9)	$\psi^{G}\ \psi^{DG}\ \psi^{Dg}$	LMAG, DIST, DIST X Species (U)	p^{G}, p^{D}	11.01	0.001
U05 (8)	$\psi^{G}\ \psi^{D}$	LMAG, DIST	$p^{G}\ p^{D} r^{D}$	11.76	0.001
U06 (8)	$\psi^{G}\ \psi^{D}$	LMAG, LMAG X Species (U), DIST	p^{G}, p^{D}	12.47	0.001
PU01 (12)	$\boldsymbol{\psi^{G}\ \psi^{D}}$	**LMAG, LMAG X Species (U), DIST, DIST X Species (U), SUBS, SUBS X Species (U)**	$\boldsymbol{p^{G}, p^{D}, r^{D}}$	**8211.03**	**0.000**
PC01 (12)	$\boldsymbol{\psi^{G}\ \psi^{DG}\ \psi^{Dg}}$	**LMAG, LMAG X Species (U), DIST, DIST X Species (C)**	$\boldsymbol{p^{G}, p^{D}, r^{D}}$	**8213.10**	**0.000**

()* = number of parameters, delta AIC = relative difference in Akaike Information Criterion, AIC_{wt} = model weight. LMAG = Link magnitude, DIST = Distance from Lake Ontario, SUBS = Percent of hard substrate, GEAR = Gear type for darters detection included in the model are seine, stationary electrofishing and boat eletrofishing. ψ^{G} Occupancy of round goby, ψ^{DG} Occupancy of a darter given round goby present, ψ^{Dg} Occupancy of a darter given round goby absent, p^{G} Detection of round goby, p^{D} Detection of darters

Table 2.6. Relative support for different formulations of a two-species occupancy model where darters occupancy was either conditional or unconditional on round goby occupancy, and where darters detection probability was either conditional or unconditional on round goby occupancy. w_i = Akaike information criterion (AIC) weight of the evidence.

Effect of round goby occupancy on darters occupancy	**Effect of round goby occupancy on darters detection**	***N***	w_i
Conditional (ψ^{G} ψ^{DG} ψ^{Dg})	Conditional (p^{G}, p^{D}, r^{D})	26	0.903
Conditional (ψ^{G} ψ^{DG} ψ^{Dg})	Unconditional (p^{G}, p^{D})	26	0.084
Unconditional (ψ^{G} ψ^{D})	Conditional (p^{G}, p^{D}, r^{D})	20	0.012
Unconditional (ψ^{G} ψ^{D})	Unconditional (p^{G}, p^{D})	20	0.001

2.3.4 Round goby and darter co-occurrence patterns

Models that assumed the occupancy of darter depended on round goby received strong support as compared to models that assumed no relationship between the two species group (conditional, $\psi^{D}\psi^{DG}\psi^{Dg}$ vs. unconditional parameterization, $\psi^{D}\psi^{DG}$ of occupancy model, Akaike weight, $w_i = 0.987$ vs 0.013, Table 2.5). From the best occupancy model (model C01, $w_i = 0.231$, Table 2.5), the mean occupancy probabilities for ψ^{D}, ψ^{DG} and ψ^{Dg} were respectively 0.59 ± 0.06, 0.90 ± 0.04 and 0.39 ± 0.06 (mean ± 1se). The best model that included hard substrate type as a covariate received little relative support (Akaike weight of 0.0038 model U01; Table 2.5). When occupancy estimation as a function of link magnitude was parameterized as unconditional and distance was parameterized as conditional, these models received highest cumulative AIC ($w_i = 0.252$, Table 2.3).

Probability of darter occupancy when sites were also occupied by round goby (ψ^{DG}) was higher than darter occupancy without round goby (ψ^{Dg}). The top two covariates indicated that darter occupancy received better support when they were estimated together ($\psi^{DG} = \psi^{Dg}$) given stream link magnitude (cumulative Akaike weight of 0.477; Table 2.3). Distance and species interaction as covariates performed better when they were estimated separately (top model covariate, AIC weight = 0.252, Table 2.3). The occupancy probabilities for round goby increased with link magnitude indicating preference for larger stream areas, although the effect was not significant (Table 2.7).

The best covariate combination indicated strongest support for the same effects of link magnitude but different effects of distance from Lake Ontario on occupancy for both round goby and darter. However, the effect of link magnitude was only significant for predicting round goby occupancy, ψ^{G} (Table 2.7). Round goby had a higher probability of occupancy at large stream than small stream magnitude size (Figure 2.2). The effect of distance from Lake Ontario shore was only significant for predicting the darter occupancy when round goby was present, ψ^{DG} (Table 2.7, Figure 2.3). Darter's probability of occupancy was lower at sites closer to Lake Ontario but higher at sites further upstreams when round goby were present (Figure 2.3).

Table 2.7. Comparison of untransformed occupancy parameter estimates of the best support model with imperfect detection probability, $p < 1$ (model C01; Table 6) and a reciprocal model that assumes detection probability, $p = 1$ (model PC01; Table 6). Estimate ± 1se. DIST = Distance from Lake Ontario; LMAG = Link magnitude.

Occupancy parameter	**Covariate parameter**	**Parameter estimates of detection probability scenario**	
		***p* < 1**	***p* = 1**
ψ^{G}	Intercept	2.33 ± 1.07*	2.04 ± 1.34
	DIST	1.09 ± 0.56	1.96 ± 0.90*
	LMAG	5.00 ± 1.83*	6.63 ± 2.81*
ψ^{DG}	Intercept	16.58 ± 7.56*	0.33 ± 0.79
	DIST	16.42 ± 7.65*	0.98 ± 0.69
	LMAG	-0.90 ± 1.04	0.18 ± 0.65
ψ^{Dg}	Intercept	-0.34 ± 1.09	0.43 ± 0.92
	DIST	3.42 ± 1.96	3.57 ± 1.95
	LMAG	-0.90 ± 1.04	0.18 ± 0.65

*Estimate is significant at 95% confidence interval

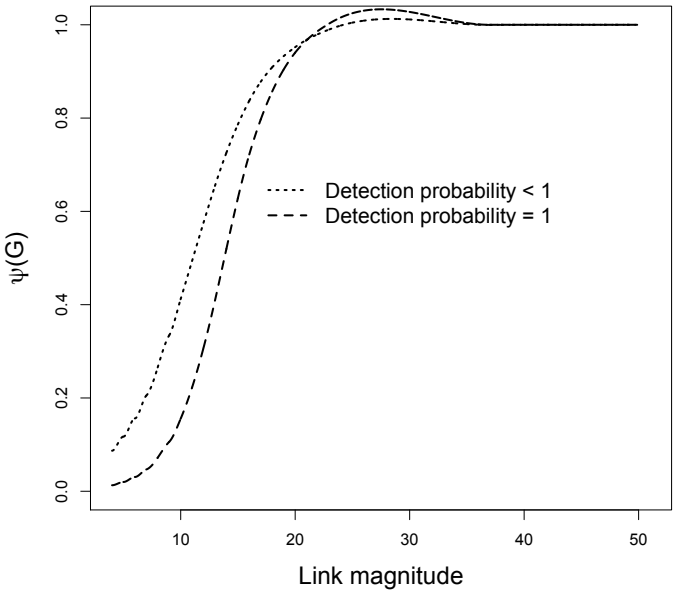


Figure 2.2. Estimated probabilities of occupancy for darters conditional on the presence or absence of round goby as a function of distance from Lake Ontario and median distance from Lake Ontario (=5.2 km) when detection probabilities, $p < 1$ and assuming perfect detection, $p = 1$. Results were obtained from model C01 and PC01.

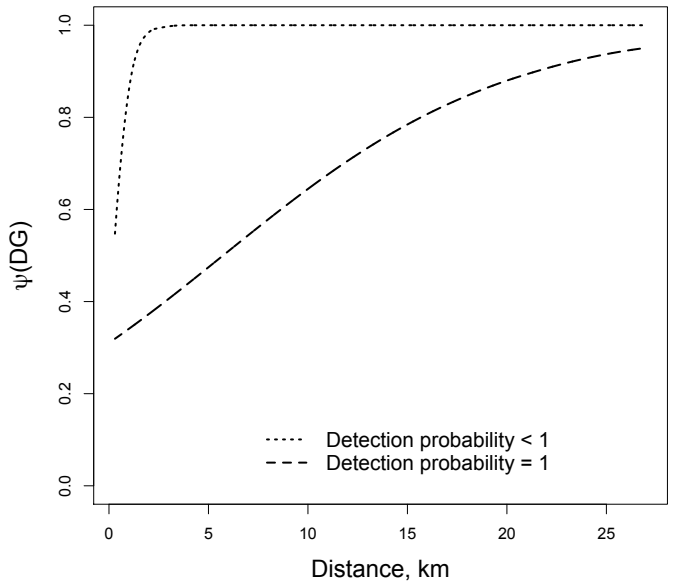


Figure 2.3. Estimated probabilities of occupancy for darters conditional on the presence or absence of round goby as a function of distance from Lake Ontario and median stream link magnitude (=10) when detection probabilities, $p < 1$ and assuming perfect detection, $p = 1$. Results were obtained from model C01 and PC01.

2.3.4 Effects of detection probabilities, $p < 1$ and $p = 1$ on species occupancy

The top model, which accounted for imperfect detection, provided much better description of the data than the model that assumed detection probability = 1 (e.g., model C01 vs model PC01; Table 2.4; Table 2.6 for parameter estimates comparison). Both detection probability scenarios ($p < 1$ vs. $p = 1$) yielded similar parameter estimates for the probability of round goby occupancy, ψ^{G}. However, by disregarding imperfect detection, models have virtually zero weight, indicating that models considering $p < 1$ provide a much better description of species co-occupancy data (Table 2.7, Figures 2.2 and 2.3). The model that assumed $p = 1$ yielded much lower occupancy probabilities (an average of 0.56 that ranged from 0.33 to 0.91) as compared to model that assumed $p < 1$ (an average of 0.90 that ranged from 0.32 to 1). The probability of occupancy for darter when round goby was not present, ψ^{Dg} and $p < 1$ yielded estimated probabilities with an average of 0.39 ranging from 0.01 to 0.99 as compared to ψ^{Dg} and $p = 1$ scenario with an average of 0.45 ranging from 0.07 to 1 (Table 2.7).

Overall, distance from the shore and link magnitude exhibited similar increasing trends with the covariate effects for occupancy parameters of ψ^{G} and ψ^{DG}. However, while the occupancy of streams by round goby decreased with link magnitude (a proxy for stream size), the probability of occupancy by round goby, ψ^{G} can also be underestimated by 2 to 15% when imperfect detection is not accounted for (Figure 2.2). Similarly, models that assumed perfect detection ($p = 1$) when round goby are present can be underestimated as much as 5 to 60% as compared to models that

assumed imperfect detection ($p < 1$) in terms of the probability of occupancy of darters, ψ^{DG} (Figure 2.3).

2.3.5 Species interaction factor

The results indicated round goby is expanding its range into upstream tributary systems and show an overlapping range with darters. The species interaction factor (SIF) was 1.71 ± 0.2 (mean ± s.e), which declined with distance from Lake Ontario and stream link magnitude. Neither round goby nor darters were more likely to co-occupy smaller streams (lower link magnitude) than sites of larger magnitude. The SIF declined with the increasing stream covariate magnitude suggesting both species could have used larger streams or farther distance from Lake Ontario independently without affecting aggregation or avoidance with each other (Figure 2.4).

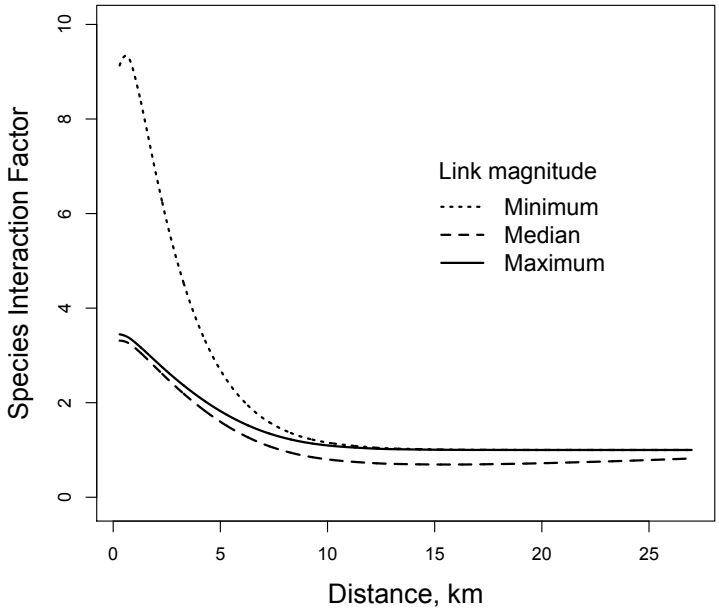


Figure 2.4. Species interaction factor as a function of distance from Lake Ontario and stream link magnitude. Results were obtained from the best support model (model C01). An SIF >1 indicates that the two species occur together more often than expected by chance, whereas an SIF <1 indicates that the two species occur together less often than expected by chance; SIF=1 indicates that the two species occur independently (no aggregation or avoidance).

2.4 Discussion

3.4.1 Species co-existence should account for imperfect detection

The ability to estimate the co-occupancy pattern of an invasive and native species is pertinent, as invasive species have been shown to pose a threat to and cause detrimental effects on native species (Balshine et al. 2005; Bergstrom and Mensinger 2009; Dubs and Corkum 1996; French III and Jude 2001). However, biological management can suffer from underestimated interactions particularly when detection probability, *p* is less than one. Accounting for imperfect detection is required for optimal co-occupancy estimation in light of species invasion. This is because the

ecological and behavioral characteristics of particular species may hamper detection. It has been suggested that because round goby and darters are relatively small and lack swim bladders (Jude et al. 1992), using a single sampling gear may be ineffective (Nett et al. 2012). As is shown in our results, estimating species occupancy without considering imperfect detection can lead to estimation of dominant species occupancy that is biased relatively low (in this case invasive species occupancy) and bias estimates of its co-occurrence rates with native fish. The 'lowered' co-occupancy estimation would essentially lead to an underestimate of the potential threats of the round goby invasion within a river system. The model comparison yielded greater average estimation of site co-occupancy that accounted for imperfect detection. Effects of invasive species on native benthic fish may occur at various temporal and spatial scales and may only be known when 'it is a little too late'. Therefore, adopting models that yield better estimates should be highlighted in order to make informed inferences about the distribution status of the species.

2.4.2 Darter distribution

We found seven darter species distributed sympatrically and yet disjunctly across our sampled area. Two darters *Etheostoma nigrum* and *E. zonale* that have been detected historically in close proximity to the study sites were not present in our study. All darters that we found in our study were listed in historical museum records for the area, providing evidence that darters collected in this study were not recent introductions. The museum records of Cornell University Museum of Vertebrates, North Carolina State Museum of Natural Sciences, Academy of Natural Sciences at

Philadelphia, University of Michigan Museum of Zoology, MCZ-Harvard University, Tulane University Museum of Natural History, Yale University Peabody Museum, California Academy of Sciences, Texas Cooperative Wildlife Collection, Royal Ontario Museum and Texas Natural History Science Center were accessed from the Fishnet2 Portal, a online collaborative institutions (www.fishnet2.net, 2012-09-26).

Given the many non-overlapping local distributions of individual native darter species, we pooled observations of individual darter species into single category. Had we not pooled these observations, at the landscape level, our occupancy model might suggest exaggerated forms of out-competition that do not account for disparate/overlapped congeneric distributions. The approach provides benefits for accounting for distribution of darter in light of the round goby invasion.

Non-overlapping congeneric darter distributions can provide evidence for existence of competition (Page and Schemske 1978). The presence of round goby, an invasive species might have caused this partial segregation by displacing darter from their habitats. However, it should be noted that failure to detect a species' presence in an occupied habitat patch could be due to sampling problems especially when the population size is small, individuals are difficult to sample, or sampling effort is limited (Gu and Swihart 2004). Furthermore, although previous studies have suggested that darter species overlap, they may differ slightly in resource, microhabitat use and/or historical origin (Sterling et al. 2012; Hlohowskyj and Wissing 1986). Because of differences in morphology, foraging method and area, intraspecific

competition among darters can experience minimal potential for resource and microhabitat competition. For instance, a study of rainbow and fantail darter interactions suggests that although the sympatric species overlapped in microhabitat and substrate uses, they did not affect each other's distribution especially during the dry season at low flow, where rainbow darter would move into deeper habitat (Schlosser and Toth 1984).

2.4.3 Co-occupancy of darter and round goby

While accounting for imperfect detection, environmental covariates are crucial to consider because landscape criteria can be a major factor in determining species distribution. Disregarding the environmental covariates can generate co-occurrence patterns that may incorrectly be interpreted as a product of competitive exclusion (Richmond et al. 2010), the situation where a dominant species, by its presence, spatially excludes a subordinate one.

We found evidence that round goby increasingly occupied sites that darter also occupied in the tributary system of North-west New York to Lake Ontario. The results corroborate previous reports that round goby populations are not just limited to the Great lakes basin within preferred rocky habitats (Clapp et al. 2001; Ray and Corkum 2001), but the populations have also expanded their range into river systems (Pennuto et al. 2010; Poos et al. 2009). The occupancy of round gobies and darters were strongly associated with distance from Lake Ontario and stream link magnitude, providing evidence on the role of water connectivity as an important component of

fish invasions. Indeed, our detection of darter was influenced by the link magnitude of the site, in which the increased presence of round goby in larger streams might affect darter detection. Such species detection interactions could arise from a density dependent effect (Stanley and Royle 2005), increasing the detection probabilities of the abundant species while reducing the detection probabilities of the sparser species. The facts that round goby prefer larger stream sizes as shown in our study reveals that stream connectivity may facilitate the dispersal of round goby. Results from the best co-occupancy model revealed water connectivity as an important mechanism by which invasive fish are spread across a riverscape (Pilliod et al. 2012), where rivers may act as limited or passive movement dispersal barrier for this species (Bronnenhuber et al. 2011). In the tributaries, the round goby populations expand when larger round gobies may induce smaller fish to leave preferred rock habitats and move to sand (Ray and Corkum 2001) and muddy habitats (Young et al. 2010) from which they disperse.

The lack support for substrate type as a surrogate for habitat resource availability in determining round goby occupancy corroborates previous reports on the flexibility of this species in its habitat use that leads to diet overlap with darters (Brush et al. 2006; Taraborelli et al. 2009). The positive occupancy association suggests overlapping habitat preferences at the patch level between round goby and darters. Although coexistence may be maintained by microhabitat or resource partitioning (Beermann and Franke 2012; Horn et al. 2012), studies particularly pertaining invasive species have also reported that coexistence of round goby with native benthic species often leads to habitat displacement of the latter species (Balshine et al. 2005; Bergstrom and

Mensinger 2009; Kipp and Ricciardi 2012). Despite the fact that both round gobies and darters are generalist feeders, round goby is more aggressive than darters (Gutowsky and Fox 2012; Marentette et al. 2012; Poulos et al. 2012). Although we found no evidence of competitive exclusion, the high cumulative AIC for conditional as compared to unconditional co-occupancy models found in this study provided strong support that indeed probability of occupancy of darter depended on the presence of round goby.

Darters were more likely to occur in smaller streams that were not occupied by round goby than in sites with round goby. Perhaps larger streams are higher in overall habitat characteristics than smaller streams that can support both round goby and darters in terms of space and resources. The finding that occupancy of round goby increased with distance from Lake Ontario provides evidence of increased invasion front of round goby into upstream tributary. Right now waterfalls and dams are known to be physical barriers in the tributaries that could deter natural dispersal of round goby from downstream to upstream areas (Kornis and Vander Zanden 2010). Upon closer inspection, large streams (link magnitude >30) occupancy probabilities for round goby were all essentially > 0.9, providing a possible explanation as to why the co-occupancy probabilities were higher even though the sites were further from Lake Ontario. These findings provided additional evidence on roles of stream networks in facilitating species invasion in tributaries (Cosentino et al. 2011; Falke et al. 2012; Kornis and Vander Zanden 2010; Peres-Neto 2004). The overlapping occurrence of round goby and darter as presented in our results suggests that round goby continues

to pose a threat to darter population in the upstream area. Adjusting to new habitats would not be a problem for round goby because they have been shown to use alternate food resources in downstream and upstream areas (Raby et al. 2010).

2.5 Conclusion

Two-species occupancy models provide a useful method for making inferences about interactions between an invasive fish, round goby, and native benthic fish, in the field. It is important to accurately document the expanding range of this invasive species so that we may provide a framework for evaluating the feasibility of control efforts in invaded waterways. Estimates of the expanding range that account for imperfect detection can be used to better predict the consequences of interspecific competition with other aquatic organisms, particularly small benthic native fish, in which round goby share similar morphological characteristics and hence are competitive for space and resource use. Connectivity of water bodies via streams results in increased probability of round goby invasion and could consequently reduce the probability of darter occurrence. These results could be used to identify and prioritize catchments and water bodies where control measures would be most effective at restoring native benthic populations. Our approach could be useful as a framework for improving investigations into questions of persistence and extirpation of native species when non-native species have already become established. We would like to reiterate that our results indicated round goby is expanding its range into upstream tributary systems and overlap with darters, and the only factor that impeded further expansion and impacts was time.

CHAPTER 3

GENETIC OCCUPANCY OF ROUND GOBY IN RELATION TO STREAM NETWORKS

Abstract

Population genetics can be used to locate barriers to dispersal by identifying genetic discontinuities that result from constrained gene flow. Here we explore how stream networks may facilitate or impair dispersal of round goby, *Neogobius melanostmus,* in the tributaries of northwestern area of New York. We tested two hypotheses about population differentiation by measuring the genetic relatedness and diversity in association with stream size and geographic position in the river system. Our results reveal evidence of limitations to the natural dispersal of round goby as shown by significant pairwise F_{ST} values among many of the local populations, admixture in Bayesian assignment analysis, and spatial patterns of genetic differentiation. We also found that goby populations in the area showed two distinct genetic clusters associated with Lake Ontario and the Erie Canal, although we did not find evidence that stream networks can constrain population mixing from where they presumably started the invasion. Lack of evidence for barriers that limit gene flow between the two clusters warrants further study that could be used to help resource managers identify areas and methods for prevention and control of invasive species.

3.1 Introduction

Understanding factors that facilitate or impede dispersal can improve our understanding of how ecological characteristics and associated communities will influence distribution and degree of ecosystem disturbance of invasive species. Through population genetics studies, researchers can locate barriers to dispersal by identifying genetic discontinuities that result from constrained gene flow. Information gathered can help resource managers to identify areas and methods for prevention and control of invasive species that are more practical and cost-effective (Klima and Travis 2012; Kornis et al. 2012).

Round goby, *Neogobius melanostomus* (Pallas, 1814), is an invasive benthic fish introduced into the Great Lakes region from Eurasia. Since its first sighting in the St. Clair River, Ontario, in the early 1990's (Jude et al. 1992) the population has expanded throughout the Great Lakes. In the associated tributaries, round goby populations have been shown to expand gradually from lakes into upstream areas (Bronnenhuber et al. 2011). Their gradual dispersal indicates there may be stream characteristics that could somehow impede the invasion processes. Indeed, although round goby exhibit a wide range of habitat requirements, studies have shown that the species prefer rocky areas over muddy habitats (Ray and Corkum 2001), and explain why certain areas are more resistant to its invasion (Cooper et al. 2007; Coulter et al. 2012). Apart from habitat preferences, round goby exhibit high-site fidelity and low dispersal during the summer (Ray and Corkum 2001; Wolfe and Marsden 1998). However, dispersal may be greater in other habitats, different life stages and other

seasons (Brownscombe and Fox 2012; Gutowsky and Fox 2012; Marentette et al. 2012). Furthermore, round goby have been shown to prefer larger than smaller streams because of greater resource availability (Ray and Corkum 2001, Phillips et al. 2003, see review in Kornis et al. 2012). Given the heterogeneous ecological properties in tributaries, those findings suggest that combined biological and ecological factors in the stream network could facilitate or constrain gene flow.

Given that the time since first introduction has been short, that natural dispersal has been high enough to propagate goby throughout the system, and that genetic differentiation takes time, it may be reasonable to assume that little to no genetic differentiation has taken place in the populations within the study area at this time. However, previous studies of round goby indicate that the species can achieve high between-site genetic differentiation in less than ten-generations, even within a relatively small spatial range (Björklund and Almqvist 2010; LaRue et al. 2011). Additionally, round goby populations in the Great Lakes Region are known to descend from multiple genetic sources originating in their native range in Eurasia liberating the species from founder effect in the invaded areas (Brown and Stepien 2008; Dillon and Stepien 2001). In the present study, I test two alternative hypotheses, not mutually exclusive, of population differentiation against the null hypothesis of no differentiation. First, if genetic differentiation stems from multiple independent introductions, then significant observable genetic differences should occur in distinct geographic regions. Second, if natural dispersal is infrequent and gene flow is limited, isolated founding events by small number of individuals should contribute to higher

rates of allele fixation through genetic drift.

To test the first hypothesis, I examined the relational ancestry of individuals collected from my study area along the southern shore of Lake Ontario, NY. If the population derives from diverse initial sources and population mixing is constrained to the geographical locations where they invaded, I expect sites to cluster genetically closer to the original source populations than to nearby geographic populations. To test the second hypothesis, I examined the relationship of genetic diversity with stream size and geographic distance position in the river system. I expect that genetic diversity should be reduced the further away the invaded sites are from the point of introduction as determined in the study due to a small number of founding individuals. Also, genetic diversity will be greater in larger streams if invasion was recent and initial populations were large. Populations were sampled at each study site, and genetic variation was analyzed at neutral molecular markers of eight microsatellite loci within the nuclear genome of round goby.

3.2 Materials and Methods

3.2.1 Selection of study sites

The study area encompasses the portion of the Lake Ontario Basin located in northwestern New York (Figure 3.1). Samples collected for this study were collected from the same study sites used for the co-occupancy estimation of round goby and darter already described (Chapter 2, this dissertation). Site selection was limited to the western portion of New York's Lake Ontario Basin, located within Hydrological Unit

(HU) code 04130001 that consists of nine 10-digit HU, as demarcated by the United States Geological Survey (USGS). Link magnitude, a measure of site position in the tributary landscape and defined by the number of un-branched source streams upstream (Shreve 1966), was calculated for each site using ArcView 3.3 and ArcGIS 9.2 (ESRI, Redlands, California). A total of 194 sites with stream link magnitudes ranging from 1 to 51 were obtained from the study area. After reconnaissance site surveys and pilot sampling, we specifically excluded sites with stream link magnitudes under 4 because our visits during the summer showed that sites below this link magnitude became dry. Within each of the 10-digit HU, we stratified site selection by randomly selecting at least two sites with stream link magnitudes 4-8, one 9-15-link stream and one above 16-link stream within a stream. This criterion was designed to ensure we sampled various stream sizes and habitat types so that the chosen sites were a relatively unbiased representation of the watershed. Of the 39 sites that were initially sampled (Chapter 2, this dissertation), I used samples from 20 sites where round goby were detected and samples were collected (Figure 3.1). Additionally, we collected samples at three sites located on Erie Canal, which may serve also as a source population for the other sites. The sites on the Erie Canal were Knowles (KNO, 43.2425, -78.3106), Lockport (LOC, 43.1848, -78.6687) and Tonawanda (TON, 43.0231, -78.8611) (site abbreviation, latitude, longitude, Figure 3.1). Approximate aquatic distances between each site, Lake Ontario, and the Erie Canal were traced and measured in Google Earth. These paths are potential routes of introduction of round goby from the two source populations.

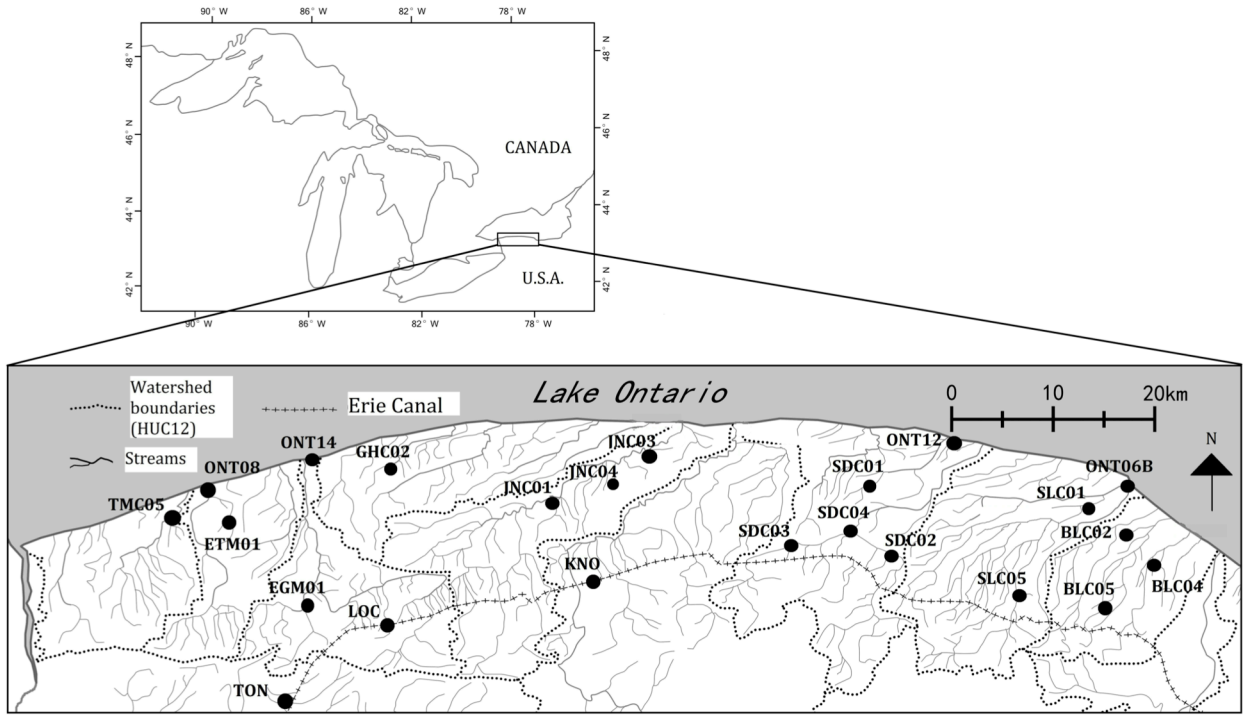


Figure 3.1. Sampled sites in northwestern New York near Lake Ontario within a U.S. Geological Survey (USGS). The sites are located within Hydrological Unit Code (HUC) 04130001. Black circles are sites where samples were used for genetic analysis.

3.2.2 Field data collection

Sampling was conducted between June and July of 2011. Samples in this study were collected using stationary electrofishing, boat electrofishing, kick-seining and angling. The selection of sampling methods depended on site conditions; we used stationary electrofishing, and kick-seining for shallow-water sites, and boat electrofishing and angling for deep-water sites. All round gobies obtained during sampling were euthanized with an overdose of MS-222 and frozen in the field prior to laboratory analysis. (For a more detailed description of sampling methods, see Appendix 2.1).

3.2.3 DNA extraction, amplification, and genotyping

A 0.5–10.0 mg clip of caudal fin tissue was pressed and air-dried to remove all ethanol from the tissue sample prior to proteinase-*K* digestion. DNA extraction followed DNeasy kit protocol by Qiagen. DNA was resuspended in 100 µL Tris–EDTA buffer (pH = 8.0). Standard polymerase chain reaction (PCR) methods were used to amplify eight polymorphic microsatellite markers (Nme3, Nme4, Nme5, Nme6, Nme7, Nme9, Nme10, and Ame133; Dufour et al. 2007, Feldheim et al. 2009). Primers were divided into three multiplexes (A, B, and C) for polymerase chain reaction (PCR) analysis, where each reaction contained basic mix of approximately 1.0 µl template, 1.0 µl BSA, 5.0 µl Qiagen mix, 0.2 µl X number of forward primers for each multiplex and 0.2 µl X number of reversed primers for each multiplex, in which RNAse-free H_2O was also added to the mix a total mix of 10 µl. Multiplex A contained primers Nme3 and Nme10 primer pairs. Multiplex B contained primers Nme4, Nme5, Nme6 and Nme9 primer pairs, whereas Multiplex C contained Nme7 and Ame133 primer pairs.

Each forward primer was fluoresced with the following dye: Fam (Invitrogen) for Nm3, Nme4, Nme6 and Nme7; Hex (Invitrogen) for Nme5, Nme10 and Ame133; Ned (Applied Biosystem) for Nme9. PCR followed the protocol of 95 °C initial denaturation for 120 s; 35 cycles of 95 °C for 15 s, and an optimized primer-specific annealing temperature (multiplex A = 48 °C; multiplex B = 55 °C; multiplex C = 59 °C) for 15 s, 72 °C extension for 30 s; followed by 72 °C final extension for 120 s and a 4 °C holding temperature ran on thermo-cycler machine. The PCR products were run on Applied BioSystems 3730xl DNA Analyzer at the Cornell Life Sciences Core Laboratories Center for visualization. Amplicons were manually analyzed and scored to size (with the base pair resolution depends on the repeat motif) using GeneMapper Software v4.0 (Applied Biosystems Inc.).

3.2.4 Standard genetic analysis prior to testing hypotheses

A number of tests were conducted to check for scoring errors and violations of assumptions prior to statistical analyses. Deviations from Hardy–Weinberg equilibrium (HWE) for each locus in each population were tested using GenePop version 3.4 with 1,000 iterations (Raymond and Rousset 1995). Linkage disequilibrium for each pair of loci in each population Arlequin version 3.5 based on the average number of alleles with 1,000 iterations (Excoffier and Lischer 2010). Sequential Bonferroni corrections for $\alpha = 0.05$ (Rice 1989) were used to evaluate the significance of deviations from HWE or linkage disequilibrium. The presence of null alleles was tested using Micro-Checker version 3.23 (Van Oosterhout et al. 2004). Finally, 10% of samples were analyzed to determine genotyping error rates, which

were calculated as the number of differently scored alleles divided by the total number of alleles across all loci (Selkoe and Toonen 2006). Observed heterozygosity, expected heterozygosity, and inbreeding coefficient (F_{IS}) were calculated for all loci in each population using GenePop version 3.4 (Raymond and Rousset 1995). Numbers of alleles per locus, allelic richness (which corrects for sample size differences among populations) were calculated using FSTAT version 2.9.3.2 (Goudet 1995).

3.2.5 Test of population differences

We used Wright's *F* statistic, also known as the fixation index, that measures the difference between the mean heterozygosity among subdivisions in a population. In Arlequin version 3.5, fixation index that measures the degree of differentiation within a population among demes (F_{ST}) further uses 'number of different alleles' as a distance method to calculate all pairs of sites' genetic differentiation (Excoffier and Lischer 2010). The significance of pairwise population comparisons was obtained with 1,000 permutations and a sequential Bonferroni for correction procedure $\alpha = 0.05$ (Rice 1989).

3.2.6 Test of population structure

Cryptic genetic structure was evaluated using Bayesian assignment analysis, implemented in Structure version 2.3.4 (Pritchard et al. 2000). Structure uses an iterative Bayesian Markov Chain Monte Carlo (MCMC) method to assign the proportion of an individual's genome to *K* distinct genetic clusters, where values of *K* can be evaluated by comparing overall likelihood scores among different values. An

admixture model with location information as a prior was used because significant pairwise F_{ST} was found among most populations (Hubisz et al. 2009). The model parameters included correlated allele frequencies, a single α, and a total run length of 10^6 generations with a burn-in period of 10^4 generations. Five replicate runs were conducted for *K* values 1–23 (23 is the total number of sampled sites) in order to evaluate likelihood scores and consistency among runs. Results generated from the simulations were collated in Structure Harvester to assess and visualize likelihood values lnP(D) of *K* (Earl and von Holdt 2012). In the program, the best model was determined based on the rate of change in the log probability of data between successive runs, otherwise known as Δ*K* method (Evanno et al. 2005). The membership coefficient matrices of multiple runs from the best *K* were aligned in CLUMPP version 1.1.2, a program that permutes all simulation replicates until they become as close a match as possible (Jakobsson and Rosenberg 2007). Estimated individual membership coefficients were visualized in Distruct version 1.1 (Rosenberg 2004). A posterior probability threshold of 0.6 was used as an individual's inferred ancestry to increase confidence in assigning individuals to the most likely genetic cluster (Coulon et al. 2008).

3.2.7 Test of effects of rivers system factors on genetic diversity

I tested the effects of link magnitude and geographic distance on population genetic of round goby by analyzing those river system attributes' relationship with their the sites' corresponding: (1) evidence of a founder effect or elevated inbreeding, (2) genetic diversity, and (3) the number of excluded individual-genotype from where they were

sampled. The predictions are: (1) as distance increases from the point of introduction, inbreeding would be elevated and genetic variation would be decreased, typical of genetic bottlenecks resulting from reduced population size due to founder effects (Nei et al. 1975; Wright 1931), and (2) larger streams would be expected to have more excluded-genotype individuals because they have higher probability of occupancy that suggest stream size determines dispersal corridor for round goby (Chapter Two, this dissertation).

Evidence of a founder effect for sites with more than ten individuals was conducted in Bottleneck version 1.2.02 (Cornuet and Luikart 1996). This software searches for the excess of heterozygotes that typically accompanies a rapid loss of alleles (Cornuet and Luikart 1996; Nei 1987). Specifically, the two-phased model (TPM) was used under the assumption that most new alleles arise by one-step mutations, with only a small percentage arising by multi-step changes (Luikart et al. 1998). This mutation model is widely accepted for analyses of microsatellite loci (Di Rienzo et al. 1994). The TPM was performed with defaults of 30 for variance and 70% stepwise mutation. 10^5 iterations of the model were used to determine expected heterozygosities in a non-bottlenecked population given the observed allelic data. A Wilcoxon sign-rank test was used to determine whether there was significant heterozygote excess for each locus relative to what would be expected in the absence of a population bottleneck.

Elevated inbreeding and reduced genetic variation, particularly with increasing distance from the point of introduction, are typical signs of genetic bottlenecks

resulting from founder effects (Nei et al. 1975; Wright 1931). In order to test for genetic bottlenecks, I conducted a series of simple regression analyses to determine if genetic diversity (sub-heading 3.2.5) of the sites within the tributaries was related to: (1) link magnitude, an index of stream size, and (2) distance from where the invasion began as inferred from population structure analysis (sub-heading 3.2.6). When a site is grouped into a population genetic cluster, the nearest river mouth associated with the cluster is regarded as where the invasion began. The distance between those points was measured using Google Earth based on the most probable distance the fish would swim.

Population differentiation may be caused by limited dispersal between sites thus constraining gene flow. Identification of the limited dispersal can be assessed through genetic assignment analysis. Genetic assignment methods have been used to draw inference about interpopulation movements i.e. where individuals were or were not born, potentially allowing direct, real-time estimates of dispersal (Paetkau et al. 2004). One of the methods can be implemented by excluding individuals, identified as immigrants by using likelihood of the individual genotype within the population where the individual has been sampled (Piry et al. 2004). I examined spatial patterns of dispersal in the study area by performing individual-level genotype exclusion in Geneclass version 2.0 (Piry et al. 2004). I used Rannala and Mountain's (1997) method with Monte Carlo resampling of 10^5 simulations (Paetkau et al. 2004) to identify fish whose genotype is excluded from sites in which they were captured, estimated with a 5% type I error threshold (Piry et al. 2004). Individuals with genotypes ‘excluded’

from the likelihood computations were then used to infer dispersal. Then I performed regression analyses to examine the relationship between (1) link magnitude and the number of excluded individuals, and (2) geographic distance from the presumable point of introduction and the number of excluded individuals. The geographic distance is the most probable distance the fish would swim between the nearest river mouths associated with the cluster is regarded as where the invasion began as measured using Google Earth. Additionally, I performed regression analyses using the same stream covariates on the number of individuals with posterior probability < 0.6 from where they were captured (sub-heading 3.2.6).

3.3 Results

3.3.1 Data set

Round goby were detected at 25 of the 39 tributary sites. Of the 25 sites where round goby was detected, samples from 20 sites in tributary and 3 sites from the Erie Canal were used for genetic analysis. A total of 559 round gobies (78 ± 0.02 cm, total length ± 1se) were genotyped at nine autosomal microsatellite loci. After genotype scoring for all the nine loci, locus Nme6 and Nme7 were discarded due to strong evidence for null alleles at the majority of sites (19/23 and 10/23 sites for Nme6 and Nme7 respectively), as indicated by Micro-Checker version 3.23 (Van Oosterhout et al. 2004). Inclusion of such loci would potentially cause 'false' deviation of heterozygote deficiency from Hardy-Weinberg proportions, biasing further population genetic analyses. Findings are therefore presented for the remaining six loci (Nme3-Nme5, Nme9, Nme10 and Ame133).

3.3.2 Genetic diversity

The majority of loci were found to be in Hardy-Weinberg equilibrium (HWE) for heterozygote deficiency after sequential Bonferroni correction, except for locus Nme5 in KNO and Ame133 in GHC02, JNC04, ONT08 and ONT12. None of the analyzed loci exhibited significant linkage disequilibrium after sequential Bonferonni corrections, with the exceptions for sites JNC04 (Nme 3/Nme 9), JNC01 (Nme 3/Nme 9) and GHC02 (Nme 10/Nme 5). The linkage disequilibrium found in these samples was likely due to the stochastic nature of population formation because no clear pattern was observed.

A total of 44 alleles were observed at the six microsatellite loci from the 23 sites. Locus Ame133 had the most alleles (11) while locus Nme4 had the least (3). The average number of alleles per locus (N_A) ranged from 3.7 (BLC05) to 5.7 (JNC03, JNC04 and BLC05) (Figure 3.2). Allelic richness (which corrects for sample size differences among populations) showed similar levels of genetic diversity across study sites, averaging 2.87–3.23 alleles per locus (Table 3.1, Figure 3.2). The majority of the sites had positive F_{IS} values, an inbreeding coefficient of samples, (range from -0.08 to 0.22), indicating that expected heterozygosity was typically higher than observed heterozygosity (Table 3.1). Average gene diversities (h) ranged from 0.58 to 0.65, and gene diversity was relatively similar among sites (visual observation of 95% CI, Table 3.1, Figure 3.2).

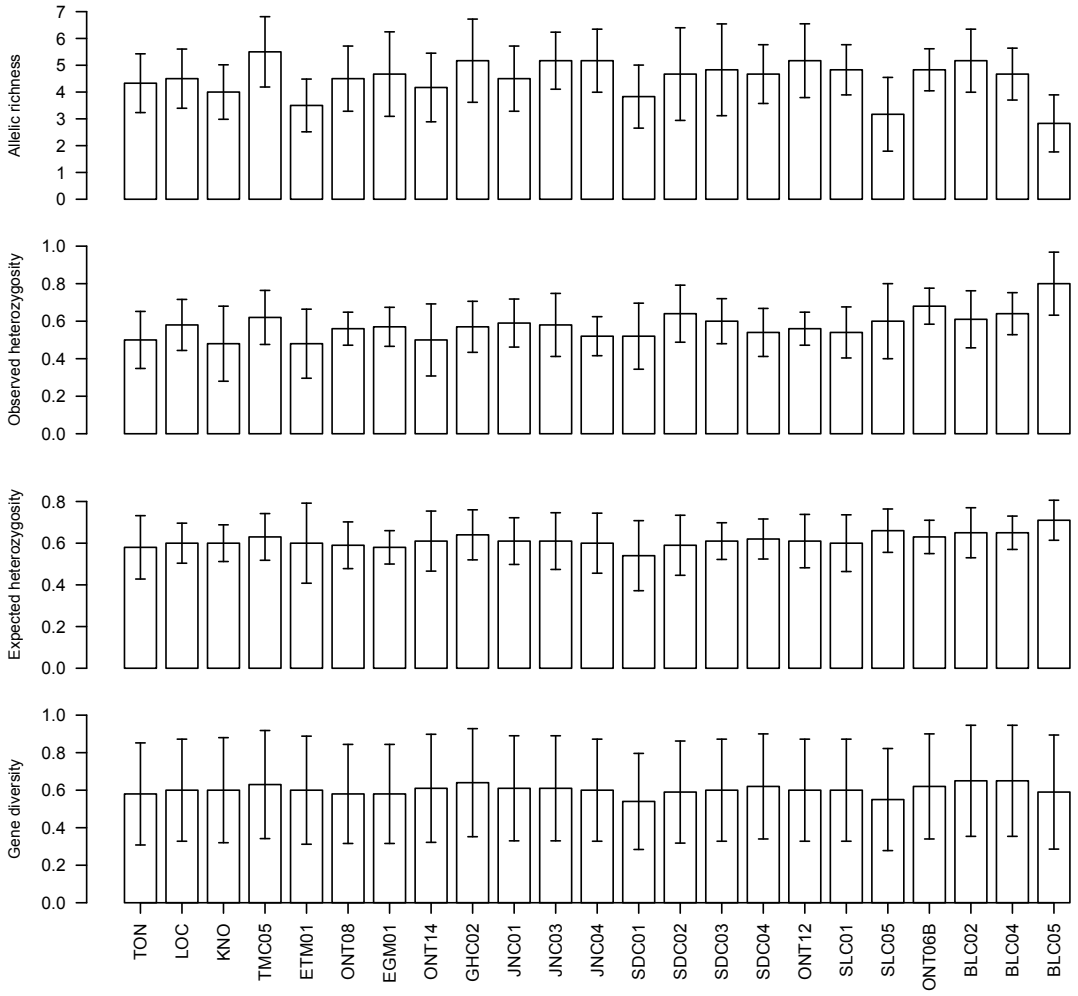


Figure 3.2. Genetic diversity before migrants identified from genotype assignment analysis were removed: number of allele N_A, observed heterozygosity H_o, expected heterozygosity H_e and gene diversity per locus from FSTAT. Bars are 95% CI. Notice that majority of sites show expected heterozygosity that was typically higher than observed heterozygosity.

Table 3.1. Site information and genetic diversity estimated for round goby: gene diversities h, average number of alleles N_A and inbreeding coefficient of sample F_{IS}, N = sample size for genetic analysis.

Waterbody: NHD HUC12	Site	N (n_1,n_2)	H_o	h	Average N_A	Allelic richness	F_{IS}
Twelvemile Creek (0213 0001 0902)	TMC05[p]	29 (3,1)	0.62	0.63	5.50	3.23	0.016
Eastbranch Creek (0213 0001 0901)	ETM01[p]	7 (0,1)	0.48	0.61	3.50	3.02	0.221
	ONT08[a]	27 (2, 0)	0.56	0.59	4.50	2.99	0.045
Eighteenmile Creek (0213 0001 0802)	EGM01[a]	30 (2,0)	0.57	0.58	4.67	2.87	0.021
	ONT14[a]	11 (2,0)	0.50	0.62	4.17	3.10	0.193
Golden Hill Creek (0213 0001 0801)	GHC02[a]	27 (3,0)	0.57	0.64	5.17	3.26	0.104
Johnston Creek (0213 0001 0704)	JNC01[p]	26 (2,1)	0.59	0.61	4.50	3.08	0.029
	JNC03[p]	31 (4,3)	0.58	0.61	5.17	3.19	0.058
	JNC04[p]	33 (2,0)	0.52	0.60	5.17	3.10	0.140
Sandy Creek (0213 0001 0603)	SDC01[p]	25 (2,0)	0.52	0.54	3.83	2.67	0.044
	SDC02[p]	24 (3,1)	0.64	0.59	4.67	3.02	-0.084
	SDC03[p]	34 (1,0)	0.60	0.61	4.83	3.01	0.018
	SDC04[p]	30 (2,0)	0.54	0.63	4.67	3.07	0.129
	ONT12[a]	34 (3,0)	0.56	0.61	5.17	3.11	0.079
Salmon Creek (0213 0001 0602)	SLC01[p]	32 (3,4)	0.54	0.60	4.83	3.10	0.103
	SLC05[p]	6 (0,0)	0.50	0.56	3.17	2.86	0.105
	ONT06B[a]	30 (2,0)	0.68	0.63	4.83	3.21	-0.072
Black Creek (0213 0001 0504)	BLC02[p]	29 (2,0)	0.61	0.65	5.17	3.36	0.063
	BLC04[p]	28 (2,0)	0.64	0.65	4.67	3.09	0.016
	BLC05[p]	4 (1,0)	0.67	0.58	4.00	2.89	-0.157
Erie Canal	KNO[n]	30 (4,11)	0.48	0.61	4.50	2.97	0.213
	LOC[n]	32 (2,0)	0.58	0.60	4.33	2.93	0.026
	TON[n]	33 (4,3)	0.50	0.58	5.50	3.23	0.144

[p]= darter present, [a]= darter absent
n_1 number of individuals with genotype excluded from the site as analyzed in Geneclass version 2.0 (Piry et al. 2004)
n_2 number of individual with posterior probability < 0.06 as analyzed in Structure version 2.3.4 (Pritchard et al. 2007; Pritchard et al. 2000).

3.3.3 Population differences

Of the 253 possible pairwise F_{ST} comparisons between sites, 111 were significant after sequential Bonferonni corrections (Table 3.2). The greatest differentiation occurred between sites ONT06B and EGM01 (pairwise F_{ST} = 0.113, p=0.00, Table 3.2, Figure 3.1). Except for BLC04-BLC02 in Blackwater Creek, none of the pairwise F_{ST} within the same stream i.e. Johnson Creek, Sandy Creek and Salmon Creek differed significantly (after sequential Bonferoni correction, Table 3.2).

Table 3.2. Pairwise F_{ST} (below diagonal) and p-value at alpha level 0.05 (above diagonal) from tests of the null hypothesis of $F_{ST} = 0$ for 23 sites in northwestern New York tributary near Lake Ontario. Bold numbers are significant F_{ST} after sequential Bonferoni correction for alpha level 0.05. Sites are grouped into population clusters (Lake and Canal) inferred from the Bayesian Markov Chain Monte Carlo (MCMC) method, which assigned the proportion of an individual's genome to two distinct genetic clusters.

	Site	N	Lake												Canal										
			KN0	**TMC05**	**ETM01**	**ONT08**	**ONT14**	**GHC02**	**JNC03**	**JNC04**	**ONT12**	**SLC01**	**ONT06B**	**BLC02**	**LOC**	**TON**	**EGM01**	**JNC01**	**SDC01**	**SDC02**	**SDC03**	**SDC04**	**SLC05**	**BLC04**	**BLC05**
Lake	**KN0**	30	--	0.00	0.05	0.00	0.07	0.05	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.13
	TMC05	29	**0.06**	--	0.55	0.00	0.59	0.11	0.20	0.15	0.01	0.00	0.00	0.23	0.00	0.07	0.00	0.01	0.00	0.01	0.00	0.00	0.22	0.00	0.12
	ETM01	7	**0.05**	0.00	--	0.25	0.61	0.39	0.41	0.26	0.27	0.30	0.10	0.36	0.51	0.70	0.05	0.18	0.20	0.24	0.04	0.35	0.68	0.03	0.82
	ONT08	27	**0.04**	**0.02**	0.02	--	0.46	0.49	0.05	0.05	0.37	0.00	0.00	0.25	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.14
	ONT14	11	0.04	0.00	0.00	0.00	--	0.92	0.77	0.92	0.83	0.14	0.00	0.47	0.01	0.23	0.00	0.01	0.03	0.08	0.01	0.00	0.17	0.00	0.30
	GHC02	27	0.02	0.01	0.01	0.00	-0.01	--	0.24	0.25	0.87	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.16
	JNC03	31	**0.03**	0.01	0.01	0.01	-0.01	0.00	--	0.92	0.02	0.01	0.00	0.08	0.01	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.24
	JNC04	33	**0.06**	0.01	0.02	0.01	-0.01	0.01	-0.01	--	0.08	0.02	0.00	0.14	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.12
	ONT12	34	**0.05**	0.02	0.01	0.00	-0.01	0.00	0.02	0.01	--	0.00	0.00	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.06
	SLC01	32	**0.08**	0.02	**0.01**	**0.03**	0.02	**0.03**	0.02	0.02	**0.04**	--	0.21	0.00	0.01	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.44	0.00	0.41
	ONT06B	30	**0.10**	**0.04**	**0.02**	**0.06**	**0.03**	**0.04**	**0.06**	**0.05**	**0.05**	0.00	--	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.05
	BLC02	29	0.04	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.00	**0.03**	**0.05**	--	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.06
Canal	**LOC**	32	**0.08**	**0.04**	**0.00**	**0.06**	0.05	**0.06**	0.05	**0.06**	**0.07**	0.03	0.03	**0.06**	--	0.42	0.01	0.15	0.03	0.11	0.01	0.31	0.59	0.01	0.35
	TON	33	0.01	0.02	-0.01	0.03	0.01	**0.03**	0.01	0.02	**0.04**	0.01	**0.04**	**0.03**	0.00	--	0.03	0.01	0.02	0.00	0.02	0.00	-0.01	0.03	-0.02
	EGM01	30	**0.09**	**0.07**	**0.05**	**0.09**	**0.07**	**0.09**	**0.06**	**0.09**	**0.11**	**0.09**	**0.11**	**0.09**	0.05	0.01	--	0.00	0.03	0.01	0.02	0.06	0.00	0.01	0.24
	JNC01	26	**0.06**	0.05	**0.02**	**0.07**	0.05	**0.06**	**0.04**	**0.05**	**0.07**	**0.05**	**0.05**	**0.07**	0.01	0.32	**0.04**	--	0.02	0.00	0.00	0.19	0.08	0.11	0.14
	SDC01	25	**0.09**	**0.06**	**0.02**	**0.08**	0.04	**0.07**	**0.05**	**0.07**	**0.08**	**0.06**	**0.07**	**0.08**	0.03	0.13	0.02	0.03	--	0.32	0.14	0.71	0.13	0.00	0.49
	SDC02	24	**0.05**	0.03	**0.00**	**0.05**	0.02	**0.04**	**0.03**	**0.05**	**0.05**	**0.05**	**0.06**	**0.04**	0.01	0.35	0.01	**0.02**	0.00	--	0.02	0.50	0.36	0.00	0.46
	SDC03	34	**0.09**	**0.06**	**0.04**	**0.09**	0.05	**0.07**	**0.07**	**0.08**	**0.09**	**0.05**	**0.06**	**0.08**	**0.02**	0.00	0.01	**0.03**	0.01	0.02	--	0.41	0.14	0.02	0.45
	SDC04	30	**0.06**	**0.04**	**0.01**	**0.06**	**0.04**	**0.05**	**0.04**	**0.06**	**0.06**	**0.04**	**0.05**	**0.06**	0.00	0.48	0.01	0.01	0.00	0.00	0.00	--	0.31	0.14	0.72
	SLC05	6	0.07	0.01	-0.01	0.03	0.03	0.04	0.02	0.04	0.05	0.00	0.03	0.03	-0.01	0.79	**0.05**	0.04	0.02	0.00	0.02	0.02	--	0.03	0.91
	BLC04	28	**0.06**	**0.05**	**0.04**	**0.09**	**0.06**	**0.06**	**0.06**	**0.08**	**0.08**	**0.07**	**0.06**	**0.08**	**0.03**	0.00	0.04	0.01	**0.04**	**0.03**	0.02	0.01	0.05	--	0.18
	BLC05	4	0.05	0.02	-0.02	0.04	0.03	0.03	0.02	0.05	0.05	0.02	0.05	0.05	0.00	0.83	0.02	0.02	-0.01	0.00	0.00	-0.02	-0.06	0.02	--

3.3.3 Population structure

The Bayesian clustering algorithm was able to assign individuals to more than one genetic cluster ($K = 2$). The majority of individuals from the sampled sites within these clusters were associated with whichever of the two major water bodies (Lake Ontario and Erie Canal, hereafter Lake and Canal) was closer (Figure 3.3). Based on a plot of estimated log likelihood for the data, the best model was $K = 4$, while a plot of the ΔK statistics of (Evanno et al. 2005) indicated $K = 2$ (Appendix 3.1). By implementing LnP(K=4) and the 0.6 threshold, individuals from site ETM01 were unassigned to any clusters, where sites and clusters association were biologically hard to interpret (Appendix 3.2). Subsequently $K = 2$ based on ΔK statistics was determined to be the best model because clusters were interpretably associated with sites and the $K = 2$ had the least variance (Appendix 2, Figure 3.3). Of 559 fish samples analyzed, 244 fish with posterior probability > 0.6 were grouped into Canal cluster, 290 fish into Lake cluster, and 25 fish were unassigned to a cluster. There were low levels of admixed individuals where only 25/559 (8.0%) individuals had a posterior probability < 0.6 to be clustered into any of the two distinct clusters in Bayesian algorithm of population structure analysis (see Table 3.1 for number of unassigned individuals).

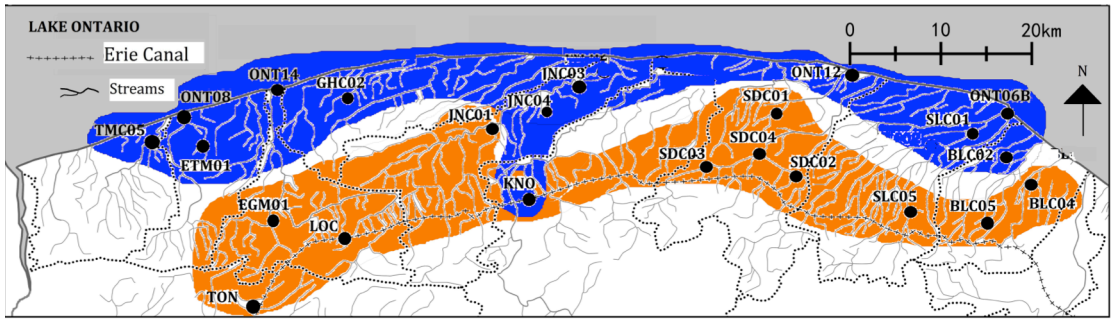
In addition, the Bayesian clustering model where K=2 yielded distinct clusters that corresponded to the geographic proximity of the sites to the Lake and Canal. For example, EGM01, JNC01, SDC01, SDC02, SDC03, SDC04, BLC04 and BLC05 – the sites that contained the most individuals with high posterior probability to one cluster,

are located near the Erie Canal, hence named Canal cluster; whereas BLC02, ONT06B, ONT12, JNC04, GHC02, ONT14 and ONT08, which contained the most individuals with high posterior probabilities to a second cluster, are located near Lake Ontario, hence named Lake cluster. Sites located between the two clusters included admixed individuals with components of both genotypes. For example, admixed individuals were increased at site SLC01, which is located between ONT06B (near Lake) and SLC05 (near Canal) (Figures 3.1 and 3.3). An important exception is KNO, which is located on the Canal, yet contained admixed individuals with higher posterior probabilities of the Lake cluster (Figures 3.1 and 3.3). Except for two sites (JNC01, BLC04); the rest of sites revealed that the clustering pattern corresponded to the relative distance to major watercourse presumably where the invasion at those sites began.

3.3.5 Genetic diversity vs. site covariates

None of our sites exhibited evidence of a founder effect, as detected by high allele frequencies and lower-than-expected heterozygosity under the TPM model of mutation after the *p*-values were adjusted using sequential Bonferroni correction (Table 3.3). No single genetic diversity measure was significantly correlated to the distance from to the presumed population source (Lake Ontario or Erie Canal) nor significantly related to link magnitude. Except for sites ETM01 and SLC05, migrant individuals were detected at all sites (Table 3.1). However, the proportion of migrant individuals per sample size at each site did not show significant relationship with link magnitude and distance. For brevity, non-significant results are not presented here.

a)



b)

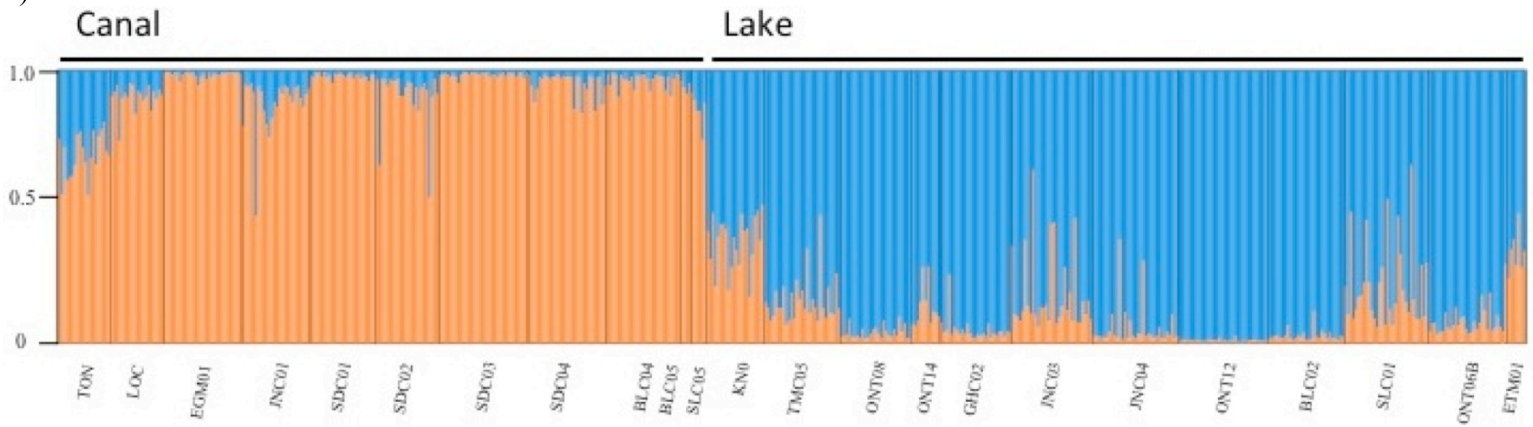


Figure 3.3. Population structure of round goby in northwestern New York: a) Sites in arbitrary blue area are of the Lake cluster, while sites outside the area are of the Canal cluster; b) Bayesian assignment results for K = 2. Each vertical line represents an individual, and colors indicate its proportion of membership to each cluster i.e. “Canal” and “Lake”.

Table 3.3. *P* value for Wilcoxon hypothesis test of one tail for heterozygote excess conducted in Bottleneck (Cornuet and Luikart 1996), for the two-phased model (TPM) for the round goby populations in northwestern New York. Individuals with posterior-probability <0.06 from the Structure Analysis were excluded from analysis. *P* values for site ETM01, SLC05 and BLC05 values are not reported because sample sizes were less than the minimum acceptable for the analysis (n=10). List of the sites are arranged based on their inferred ancestry cluster and inbreeding coefficient, F_{IS}.

Site	Cluster*	TPM	F_{IS}
TMC05	Lake	0.023[ns]	0.009
ONT08	Lake	0.039 [ns]	0.0451
ONT14	Lake	0.039 [ns]	0.1932
GHC02	Lake	0.055	0.1036
JNC03	Lake	0.039 [ns]	0.0714
JNC04	Lake	0.078	0.1404
ONT12	Lake	0.016 [ns]	0.079
SLC01	Lake	0.016 [ns]	0.0867
ONT06B	Lake	0.016 [ns]	-0.0718
BLC02	Lake	0.008 [ns]	0.0627
KNO	Lake	0.016 [ns]	0.247
BLC04	Erie	0.039 [ns]	0.016
LOC	Erie	0.344	0.026
TON	Erie	0.039 [ns]	0.156
EGM01	Erie	0.078	0.021
JNC01	Erie	0.055	0.006
SDC01	Erie	0.008 [ns]	0.044
SDC02	Erie	0.219	-0.115
SDC03	Erie	0.039 [ns]	0.018
SDC04	Erie	0.055	0.129

[n] Percent of individuals with posterior-probability the threshold > 0.6 were retained in the site groups they were sampled.
[ns] not significant after Bonferoni correction for alpha = 0.05.

3.4 Discussion

In the present study we found that round goby populations showed high within-site genetic diversity and significant local scale genetic differentiations among study sites, corroborating the studies conducted along Lake Michigan and in the Baltic Sea (Björklund and Almqvist 2010, LaRue et al. 2011)**.** Multiple genetic sources due to multiple introductions have been identified as one of the major causes that contribute to the successful proliferation of round goby invasion across the Great Lakes region (Brown and Stepien 2009; Stepien et al. 2005; Stepien and Tumeo 2006), liberating the species from founder effects of loss genetic diversity that are typically experienced by small population sizes of introduced species (Chen et al. 2012; Dlugosch and Parker 2008; Klima and Travis 2012; Roman and Darling 2007; Zhan et al. 2012).

Our findings confirm those of two earlier studies conducted in other regions that demonstrate the same pattern of genetic differentiation in sub-populations of round goby despite its relatively recent introduction to these landscapes (see Björklund and Almqvist 2010, LaRue et al. 2011). LaRue et al. (2011) attributed the genetic differentiation to limited natural dispersal with frequent long-distance dispersal through anthropogenic activities such as commercial shipping. Björklund and Almqvist (2010) found that as round goby occupy different habitats along the Polish and Latvian shorelines, this might limit their genetic homogenization. Our study of the population genetic structure also suggests a low degree of panmixis i.e. low gene flow within and between watersheds in northwestern New York. These results are supported by the presence of many significant F_{ST} between our sample sites (121 of

252 pairwise F_{ST} were significant).

3.4.1 Population structure

Apart from Lake Ontario, we also found that goby populations also invaded northwest New York's tributaries from the Erie Canal as indicated by the distinct population structure that coincides with the location of our study sites as well as genetic admixture between those two major watercourses. F_{ST} for some site pairs of adjacent distance were more significantly different than site pairs that were farther distance.

Despite the suggestion of limited gene flow between the two most probable points of invasion, Lake Ontario and Erie Canal, as evidenced by the presence of two population clusters, we found admixed individuals at the KNO site on Erie Canal, where individuals collected from this site had the probability of almost equally descending from Lake and Canal groups. Genetic differentiation is generally related to geographic distance (i.e. isolation by distance), but can be periodically interrupted due to gene flow (LaRue et al. 2011). For example, there were three sites sampled in Blackwater Creek, BLC02, BLC04 and BLC05. The latter two (BLC04 and BLC05) are located closer to Erie Canal than BLC02. Indeed, Bayesian clustering showed that BLC04 and BLC05 consisted of individuals grouped into the Canal cluster, while one (BLC02) consisted of individuals grouped into the Lake cluster. The fact that BLC05 were significantly different from all sites in the study perhaps due in part to small population size of four individuals. The spatial patterns of genetic differentiation evident in the Bayesian assignment suggest that populations generally experience

restricted gene flow, clustered into Lake and Canal populations. Björklund and Almqvist (2010) suggest the high rate of genetic differentiation in their study area could be due to low effective population size, as a result of a system where only a few parents have, by chance, a large number of surviving fry ("The Hedgehock effect", Waples 1998).

3.4.2 Stream network and population mixing

Despite evidence of limited dispersal, our study provides little evidence that genetic variation decreases with distance from the presumed source populations or as stream size decreases, as would be expected due to founder effects (Nei et al. 1975; Wright 1931). One possible explanation is that our study examines the invasion process at too small a scale to infer rates of gene flow without conducting temporal replicates. Furthermore, selection of random sites rather than systematic sampling points along the same streams as in Bronnenhuber et al. (2011) and LaRue et al. (2011), necessitated cross-stream comparisons rather than measuring gene flow along 'actual' invasion corridors.

Population differentiation is expected if a limited number of migrants are exchanged among populations; nearby populations are more likely to exchange migrants than populations located further apart, resulting in genetic similarities among nearby sites (LaRue et al. 2011). Such conditions are particularly true for fish invading tributary systems because stream networks can promote or restrict gene flow (Williams et al. 2003). Paradoxically population differentiation is only possible if round goby migrate

at the rather modest range of 1 km/year (Bergstrom et al. 2008) and the rate of larval dispersal among populations is low. We observed more significant between-site differentiation in the Lake than within the Canal cluster, whereas the inbreeding coefficient, $F_{IS,}$ was higher in the Lake than the Canal cluster. This indicates that there is greater isolation between populations in the Lake Ontario populations than in the Erie Canal populations, despite round goby are habitat generalists. However, one recent study suggests that round goby can disperse quickly along streams, approximately 9 km in a period of 8 months (Brownscombe and Fox 2012), although such rapid dispersal also depends on the attributes of the stream (e.g. water velocity, how much contiguous habitat is available, seasonal hydrology, available food resources, fish community resistance). This rapid dispersal would tend to unify populations along the same stream. Populations from the disparate ancestry might have invaded the area through a different 'source' water-body as seen in the Bayesian cluster patterns in this study (Lake vs. Canal cluster).

Therefore, we did not find evidence that stream networks can constrain population mixing. Although the mechanisms for genetic unification and differentiation include short-distance dispersal to nearby sites and longer-distance dispersal to remote sites (Bronnenhuber et al. 2011; LaRue et al. 2011), tributaries, dams and waterfalls may limit dispersal of round goby upstream (Kornis and Vander Zanden 2010). These short- and longer-distance dispersal classes may be caused by different mechanisms; short-distance dispersal may be natural, while long-distance dispersal may be anthropogenic (e.g., commercial shipping or bait bucket transfers) (Bronnenhuber et

al. 2011; Hensler and Jude 2007; LaRue et al. 2011; Lynch and Mensinger 2012).

3.5 Conclusion

Population genetic differentiation of round goby in its invaded range was detected at fine scales, where each site did not appear to exhibit apparent founder effects. Individual round gobies in the study area can be assigned to at least two distinct lineage clusters, and these predictably coincide with the closest major water bodies, which presumably served as source populations. Genetic discontinuities may reflect barriers to dispersal or high site fidelity, as one would expect from fish with lack swim bladders. When non-significant effects of stream network provide lack evidence of limited dispersal barriers, examining population structure provided insights into the mechanisms of invasion in the area. Understanding patterns of gene flow in areas where round goby have already become established allows the development of wise management measures, where they are continuing to invade.

APPENDICES

Appendix 1.1. Pairwise geographic distance and Mahalanobis distance between locations.

Location pair		Habitat pair	Geographic distance, km	Mahalanobis distance	
				Round goby	Pumpkinseed sunfish
LONT01	LONT02	Lake-Stream	81	3.4149	2.7312
LONT01	LONT03	Lake-Lake	98	3.0513	2.2564
LONT01	LONT04	Lake-Lake	115	2.2779	1.3373
LONT01	LONT05	Lake-Lake	171	2.3691	1.9265
LONT01	LONT06	Lake-Stream	235	2.0115	2.1916
LONT01	LONT07	Lake-Stream	273	2.7823	3.1578
LONT01	LONT08	Lake-Stream	349	2.6992	2.6284
LONT02	LONT03	Lake-Stream	17	1.8704	2.9487
LONT02	LONT04	Lake-Stream	34	2.4983	2.3025
LONT02	LONT05	Lake-Stream	90	3.4493	2.4862
LONT02	LONT06	Stream-Stream	154	3.5827	3.0537
LONT02	LONT07	Stream-Stream	192	2.4403	3.4182
LONT02	LONT08	Stream-Stream	268	3.5303	3.2787
LONT03	LONT04	Lake-Lake	17	1.907	1.7657
LONT03	LONT05	Lake-Lake	73	3.0193	1.5733
LONT03	LONT06	Lake-Stream	137	2.826	2.0384
LONT03	LONT07	Lake-Stream	175	1.7762	2.3622
LONT03	LONT08	Lake-Stream	251	2.5338	2.6844
LONT04	LONT05	Lake-Lake	56	2.5144	1.8168
LONT04	LONT06	Lake-Stream	120	1.952	1.8694
LONT04	LONT07	Lake-Stream	158	2.2542	2.5596
LONT04	LONT08	Lake-Stream	234	2.7903	2.4857
LONT05	LONT06	Lake-Stream	64	2.8369	1.8799
LONT05	LONT07	Lake-Stream	102	2.7146	2.3907
LONT05	LONT08	Lake-Stream	178	2.9392	2.3225
LONT06	LONT07	Stream-Stream	38	2.5364	1.7633
LONT06	LONT08	Stream-Stream	114	3.0371	1.0749
LONT07	LONT08	Stream-Stream	76	2.2975	1.739

Appendix 2.1. Procedures for each sampling gear

Kickseining

A seine net about 1.8-m tall and 3-m width with 5-mm mesh size was used to sample in shallow water sites in a systematic zigzag fashion. Efforts for each site sampled with this gear covered a surface area of average 320-m^2 (approximately 80-m length, 4-m width). We sampled an average of 27 quadrats of kick-seining (minimum 24, maximum 32) moving zigzag from downstream to upstream of the site. At each quadrat of about 6-m^2 (2-m X 3-m), two individuals held the seine net at the ends and two individuals kicked the sediments in order to guide fish to downstream into the seine. We conducted seine hauls in up-stream direction at areas where kicking was not possible (due to fast or deep water that impeded kicking).

Stationary electro-fishing

For electrofishing, a 3,500-W generator and a transformer—DC voltage output controller of 125–250 V were installed on stream banks. One end of a cathode array with 4-m steel cables was connected to the transformer and the other end was centrally positioned in each of the sites. One end of a 100-m cable was connected to the transformer and another end of the cable was connected to a handheld steel ring (0.5-m diameter) anode. Using this configuration, the power supply generated a strong electric current across the sampling area where the field crew would maneuver through the sampling station, putting the current into the water to stun the fish. Approximately 60 to 100 m of channel length was sampled in a zigzag fashion from downstream to upstream with a single pass. Electrofishing halted intermittently after approximately 2 minutes or 10 m of channel length of electrofishing and the transformer repositioned.

Boat electrofishing

Sampling was conducted by using an electrofishing boat for approximately 1.5 hours of total effort for each locality. The boat was equipped with a Smith-Root Type VI-A transformer, a Honda 5000 watt generator and two protruding bars with anode droppers suspended in front of the boat, where the boat hull served as the cathode array. Within each locality, sampling effort was relatively equally distributed on each available microhabitat within approximately 100m radius of the boat launching area. We electrofished by putting direct current (DC) into the water (the current was controlled between 25-250V) for about 2-3 minutes at any given microhabitats. Stunned fish between the electrodes were caught using dip nets. Boat driving speed was matched to the average walking speed of 4.8 km per hour used by the walking field crew in shallow water sites (Singkran and

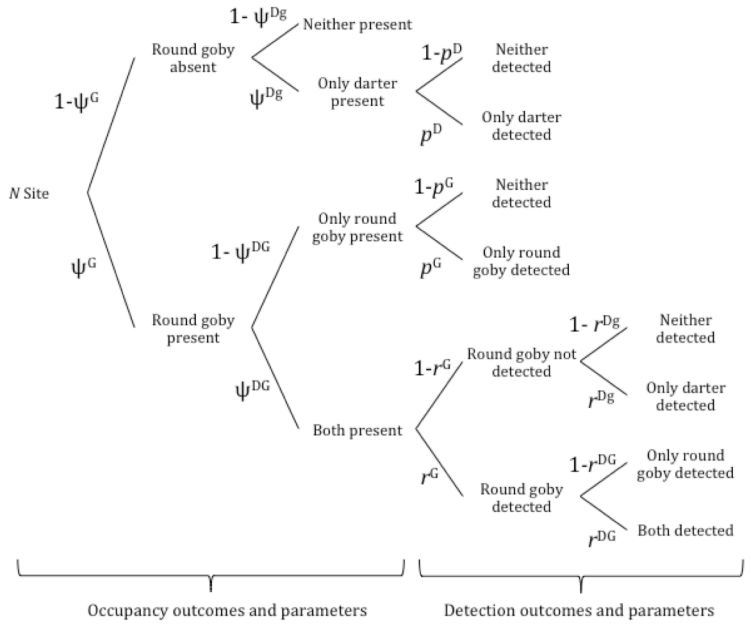
Meixler 2008).

Angling

A four-angler effort was used to catch round goby at each site. Anglers were relatively equally spaced over the stream banks while angling. We used Shimano angling poles, size 9 to 10 hook, monofilament line baited with worms and angling halted after approximately one hour. Angled gobies were kept in a water bucket and removed from the sampling area.

Appendix 2.2. a) Model structure for the conditional parameterization of the single-season two-species occupancy model; b) Description of corresponding notations used in the model structure.

a)



b)

Parameter	Description
ψ^{G}	Probability of occupancy for round goby
ψ^{DG}	Probability of occupancy for at least one darter species, given round goby are present
ψ^{Dg}	Probability of occupancy for at least one darter species, given round goby are absent
p^{G}	Probability of detection for round goby, given all darter are absent
p^{D}	Probability of detection for at least one darter species, given round goby is absent
r^{G}	Probability of detection for round goby, given both round goby and a darter species are present
r^{DG}	Probability of detection for at least one darter species, given both round goby and at least one darter species are present and round goby is detected
r^{Dg}	Probability of detection for at least one darter species, given both round goby and at least one darter species are present and round goby is not detected

Appendix 2.3. Detection model description

For each site survey in this study, the notation for detection history would be $_i1/0_i1/0$, where the subscript *i* and 1 or 0 are the site number with observation histories of the respective two sampling methods used for each site. For example, if round goby was detected using kick-seine and stationary electrofisihing at a site but darter was only detected using stationary electrofising at the same site, the detection history would be $_i1_i1$ for round goby and $_i1_i0$ for darter, defined as $\mathbf{p}_i^{\{11\},\{01\}}$. The probability of location *i* being in each of the four respective occupancy states becomes the elements of Φ and sum to 1.

I used a conditional parameterization of the two-species occupancy model to test the occurrence pattern of round gobies and darter in the study area (Richmond et al. 2010). The model estimates the probability of occupancy for a subordinate species conditional upon the presence of a dominant species (Richmond et al. 2010). The two-species occupancy model, which utilizes maximum likelihood theory, consists of an unconditional probability of occupancy parameter for the dominant species and a conditional probability of occupancy for subordinate species in relation to the presence or absence of dominant species.

Each of the occupancy states consists of probabilities of detection for round goby and darter that are conditional on the occupancy status of both species (Appendix of model hierarchical structure). Consider the observation history example I provided above. The detection observations can reflect an occupancy state where both species being present at the site (round goby detected and darter undetected by kick-seine; and both detected by stationary electro-fishing). Therefore for site, the probability of observing the given detection histories for each of the four conditional occupancy states can be written as the following vector column,

$$\mathbf{p}_i^{\{11\},\{01\}} = \begin{bmatrix} r_{i1}^{G}\left(1 - r_{i1}^{DG}\right) r_{i2}^{DG} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where r^{G} is the probability of detecting round goby when both species are present, r^{DG} is the probability of detecting darter when both species are present and round goby is detected (Table 1, Figure 2), *i* corresponds to the location, and the subscript number corresponds to the gear type. Because both round goby and darter were present at the site, only one probability of detection for conditional occupancy when both species are present (ψ^{DG}) is calculated in the vector. The probability of observing the given detection histories for the other three possible occupancy states [only round goby was present ($1-\psi^{DG}$); only darter species were present (ψ^{Dg}); neither species is present ($1-\psi^{Dg}$)] is essentially 0, because round goby and darter were detected, and the model, as parameterized, does not account for false presences.

Appendix 2.4. Determination of the best detection covariate for round goby and darter using single-season single-species occupancy model

Following the methods to reduce the number models for occupancy model by Richmond et al. (2010), I determined the best detection covariate for both species by using single-season single species occupancy model.

To initially identify the best detection model (p), I held occupancy constant ($\psi(.)$) while fitting four single-season single-species occupancy models (MacKenzie et al. 2002). The model separately calculates the probability of detecting an individual given a site and the probability that at least one individual is present at the site given a time frame using maximum likelihood theory (MacKenzie et al. 2002). I tested the best detection model for round goby and darter to test the effects of variable types of gear used in this study (seine, stationary electrofishing, boat electrofishing and angling) and stream size, represented by link magnitude. For this single-season single-species detection model for darter, I did not include anglers as an observation covariate because darter was not detected at any sites using this method (these observation were treated as missing values). I hypothesized that the detection would decrease as stream link magnitude increases, because both round goby and darter detection could be impeded by larger stream sizes (where link magnitude was used as a covariate to test this hypothesis).

Effects of covariates on round goby and darter detection probability

Factors receiving the most support as affecting detection rate were the same for darters and round goby (Tables 4 and 5). The best-supported detection model was the model that included stream link magnitude as a detection covariate. The best-supported detection model for round goby had an Akaike weight of 0.9992, while 0.7471 for darter. Average detection probability calculated from the best round goby detection model was the highest and perfect with electrofishing (1.0), followed by angling (an average of 0.85, ranged from 0.63 to 0.98), stationary electrofishing (an average of 0.43, ranged from 0.26 to 0.93) and seining (an average of 0.25, ranged from 0.21 to 0.89). For darter, the average detection probability calculated from the best darter detection model was the highest with seining (an average of 0.72, ranging from 0.63 to 0.91), followed by stationary electrofishing (an average of 0.61, ranging from 0.51 to 0.88) and boat eletrofishing (an average of 0.33, ranging from 0.19 to 0.51).

Appendix 2.4. cont.

Detection models for darters with constant occupancy. GEAR^{æ}= *Gear used for sampling and included in the model for darters were seine, stationary electrofishing and boat electrofishing;* LMAG = *Stream link magnitude;* AIC = *Akaike information criterion.* Ψ = *probability of occupancy, p = probability of detection. ΔAIC = the difference between the Akaike information criterion (*AIC*) of model of interest and the best-supported model.* w_i = *AIC weight of the evidence.*

Model	ΔAIC	w_i	Model likelihood	*K*	-2*log-likelihood
ψ(.);p(LMAG)	0.00	0.9992	1.0000	6	76.33
ψ(.);p(.)	0.31	0.0008	0.0008	5	92.64
ψ(.);p(GEAR^{æ})	0.97	0.0000	0.0000	8	92.64

Detection models for round goby with constant occupancy. GEAR^{ŧ} = *Gear used for sampling and included in the model for darters were seine, stationary electrofishing, boat electrofishing and angling;* LMAG = *Stream link magnitude;* AIC = *Akaike information criterion.* Ψ = *probability of occupancy, p = probability of detection.* ΔAIC = *the difference between the Akaike information criterion (*AIC*) of model of interest and the best-supported model.* w_i = *weight of the evidence.* w_i=AIC *weight of the evidence.*

Model	ΔAIC	w_i	Model likelihood	*K*	-2*log-likelihood
ψ(.);p(LMAG)	0.00	0.7471	1.0000	5	78.76
ψ(.);p(.)	2.42	0.2228	0.2982	4	83.18
ψ(.);p(GEAR)	6.42	0.0301	0.0404	6	83.18

Appendix 2.5. Model candidates. DIST = distance from Lake Ontario, km; LMAG = stream link magnitude; SUBS = hard substrate type, %. Ψ = probability of occupancy, p = probability of detection. ΔAIC = the difference between the Akaike information criterion (AIC) of model of interest and the best-supported model. w_i = AIC weight of the evidence. Occupancy code: A= conditional, B = unconditional.

Occupancy	Detection	AIC	ΔAIC	AIC	w_i	No. Parameter	Model Likelihood
A_15_lmag,dist,distXspecies	(Cond),lmagXspecies(Uncond)_C	144.54	0	0.231	1	12	120.54
A_14_lmag,dist,distXspecies	(Uncond),lmagXspecies(Uncond)_C	144.78	0.24	0.2049	0.8869	11	122.78
A_13_lmag,dist,distXspecies	(Cond),lmagXspecies(Cond)_C	144.84	0.3	0.1988	0.8607	13	118.84
A_12_lmag,dist,distXspecies	(Uncond),lmagXspecies(Cond)_C	144.86	0.32	0.1968	0.8521	12	120.86
A_09_lmag,dist,lmagXspecies	(Uncond)_C	147.09	2.55	0.0645	0.2794	10	127.09
A_15_lmag,dist,distXspecies	(Cond),lmagXspecies(Uncond)	149.35	4.81	0.0209	0.0903	11	127.35
A_14_lmag,dist,distXspecies	(Uncond),lmagXspecies(Uncond)	149.47	4.93	0.0196	0.085	10	129.47
A_12_lmag,dist,distXspecies	(Uncond),lmagXspecies(Cond)	149.74	5.2	0.0172	0.0743	11	127.74
A_13_lmag,dist,distXspecies	(Cond),lmagXspecies(Cond)	149.75	5.21	0.0171	0.0739	12	125.75
A_09_lmag,dist,lmagXspecies	(Uncond),	152.18	7.64	0.0051	0.0219	9	134.18
A_08_lmag,dist,lmagXspecies	(Cond)_C	152.38	7.84	0.0046	0.0198	7	138.38
B_19_lmag,dist,subs,distXspecies	(Uncond),subsXspecies(Uncond),lmagXspecies(Uncond)_C	152.78	8.24	0.0038	0.0162	12	128.78
B_06_lmag,dist,lmagXspecies	(Uncond),_C	152.86	8.32	0.0036	0.0156	9	134.86
A_08_lmag,dist,lmagXspecies	(Cond),	153.72	9.18	0.0023	0.0102	10	133.72
B_08_lmag,dist,distXspecies	(Uncond),lmagXspecies(Uncond)_C	153.87	9.33	0.0022	0.0094	10	133.87
B_05_lmag,dist,distXspecies	(Uncond),_C	154.31	9.77	0.0017	0.0076	9	136.31
A_55_lmag,dist,subs,distXspecies	(Cond),subsXspecies(Cond),lmagXspecies(Cond)_C	154.43	9.89	0.0016	0.0071	13	128.43
A_06_lmag,dist,distXspecies	(Uncond),	155.55	11.01	0.0009	0.0041	9	137.55
B_04_lmag,dist,_C	Constant	156.3	11.76	0.0006	0.0028	8	140.3
B_06_lmag,dist,lmagXspecies	(Uncond),	157.01	12.47	0.0005	0.002	8	141.01
B_19_lmag,dist,subs,distXspecies	(Uncond),subsXspecies(Uncond),lmagXspecies(Uncond)	157.11	12.57	0.0004	0.0019	11	135.11
A_07_lmag,dist,distXspecies	(Cond),	157.55	13.01	0.0003	0.0015	10	137.55
B_08_lmag,dist,distXspecies	(Uncond),lmagXspecies(Uncond)	158.31	13.77	0.0002	0.001	9	140.31
A_28_lmag,subs,_C	Constant	158.4	13.86	0.0002	0.001	9	140.4
A_05_lmag,dist,	Constant	158.62	14.08	0.0002	0.0009	8	142.62
A_30_lmag,subs,lmagXspecies	(Uncond),_C	158.95	14.41	0.0002	0.0007	10	138.95
B_05_lmag,dist,distXspecies	(Uncond),	159.1	14.56	0.0002	0.0007	8	143.1
A_55_lmag,dist,subs,distXspecies	(Cond),subsXspecies(Cond),lmagXspecies(Cond)	160.85	16.31	0.0001	0.0003	12	136.85
A_29_lmag,subs,lmagXspecies	(Cond),_C	160.95	16.41	0.0001	0.0003	11	138.95
B_18_lmag,subs,subsXspeceis	(Uncond),lmagXspecies(Uncond)_C	160.98	16.44	0.0001	0.0003	10	140.98
B_04_lmag,dist,	Constant	161.3	16.76	0.0001	0.0002	7	147.3
B_12_lmag,subs,_C	Constant	163.25	18.71	0	0.0001	8	147.25
B_07_lmag,lmagXspecies	(Uncond),_C	163.34	18.8	0	0.0001	8	147.34
A_04_lmag_C	Constant	163.54	19	0	0.0001	7	149.54
A_35_subs,subsXspeceis	(Uncond),_C	163.62	19.08	0	0.0001	9	145.62

Appendix 2.5 cont'

Occupancy	Detection	AIC	ΔAIC	AIC	w_i	No. Parameter	Model Likelihood
B_18_lmag,subs,subsXspeceis	(Uncond),lmagXspecies(Uncond)	163.92	19.38	0	0.0001	9	145.92
A_28_lmag,subs,	Constant	164.31	19.77	0	0.0001	8	148.31
A_10_lmag,lmagXspecies	(Cond)_C	164.32	19.78	0	0.0001	7	150.32
A_11_lmag,lmagXspecies	(Uncond)_C	164.32	19.78	0	0.0001	7	150.32
B_13_lmag,subs,lmagXspecies	(Uncond),_C	164.45	19.91	0	0	9	146.45
A_10_lmag,lmagXspecies	(Cond),	164.8	20.26	0	0	9	146.8
A_11_lmag,lmagXspecies	(Uncond),	164.8	20.26	0	0	9	146.8
A_34_subs,subsXspeceis	(Cond),_C	164.99	20.45	0	0	10	144.99
B_14_lmag,subs,subsXspeceis	(Uncond),_C	164.99	20.45	0	0	9	146.99
A_30_lmag,subs,lmagXspecies	(Uncond),	165.11	20.57	0	0	9	147.11
B_03_lmag,_C	Constant	165.75	21.21	0	0	7	151.75
B_07_lmag,lmagXspecies	(Uncond),	166.52	21.98	0	0	7	152.52
A_18_dist,subs,distXspecies	(Uncond),_C	166.68	22.14	0	0	10	146.68
A_04_lmag,	Constant	167.07	22.53	0	0	7	153.07
A_29_lmag,subs,lmagXspecies	(Cond),	167.11	22.57	0	0	10	147.11
A_07_lmag,dist,distXspecies	(Cond)_C	167.28	22.74	0	0	7	153.28
A_05_lmag,dist_C	Constant	167.28	22.74	0	0	7	153.28
A_06_lmag,dist,distXspecies	(Uncond)_C	167.28	22.74	0	0	7	153.28
B_12_lmag,subs,	Constant	167.71	23.17	0	0	7	153.71
B_14_lmag,subs,subsXspeceis	(Uncond),	168.03	23.49	0	0	8	152.03
A_35_subs,subsXspeceis	(Uncond),	168.11	23.57	0	0	8	152.11
A_17_dist,subs,distXspecies	(Cond),_C	168.47	23.93	0	0	11	146.47
B_13_lmag,subs,lmagXspecies	(Uncond),	169.01	24.47	0	0	8	153.01
A_16_dist,subs,_C	Constant	169.21	24.67	0	0	9	151.21
B_03_lmag,	Constant	169.45	24.91	0	0	6	157.45
A_34_subs,subsXspeceis	(Cond),	169.47	24.93	0	0	9	151.47
A_02_dist,distXspecies	(Uncond)_C	170.54	26	0	0	7	156.54
A_03_dist,distXspecies	(Cond),_C	170.54	26	0	0	7	156.54
A_03_dist,distXspecies	(Cond),	171.01	26.47	0	0	9	153.01
A_02_dist,distXspecies	(Uncond),	171.06	26.52	0	0	8	155.06
A_33_subs,_C	Constant	171.47	26.93	0	0	8	155.47
B_11_dist,subs,subsXspeceis	(Uncond),_C	172.02	27.48	0	0	9	154.02
B_17_dist,subs,distXspecies	(Uncond),subsXspecies(Uncond)_C	172.68	28.14	0	0	10	152.68
A_18_dist,subs,distXspecies	(Uncond),	173.04	28.5	0	0	9	155.04
A_17_dist,subs,distXspecies	(Cond),	173.99	29.45	0	0	10	153.99
A_01_dist_C	Constant	174.08	29.54	0	0	7	160.08
B_16_subs,subsXspeceis	(Uncond),_C	174.42	29.88	0	0	8	158.42
B_11_dist,subs,subsXspeceis	(Uncond),	175.24	30.7	0	0	8	159.24
A_16_dist,subs,	Constant	175.55	31.01	0	0	8	159.55
A_psi(.)_C	Constant	175.55	31.01	0	0	7	161.55

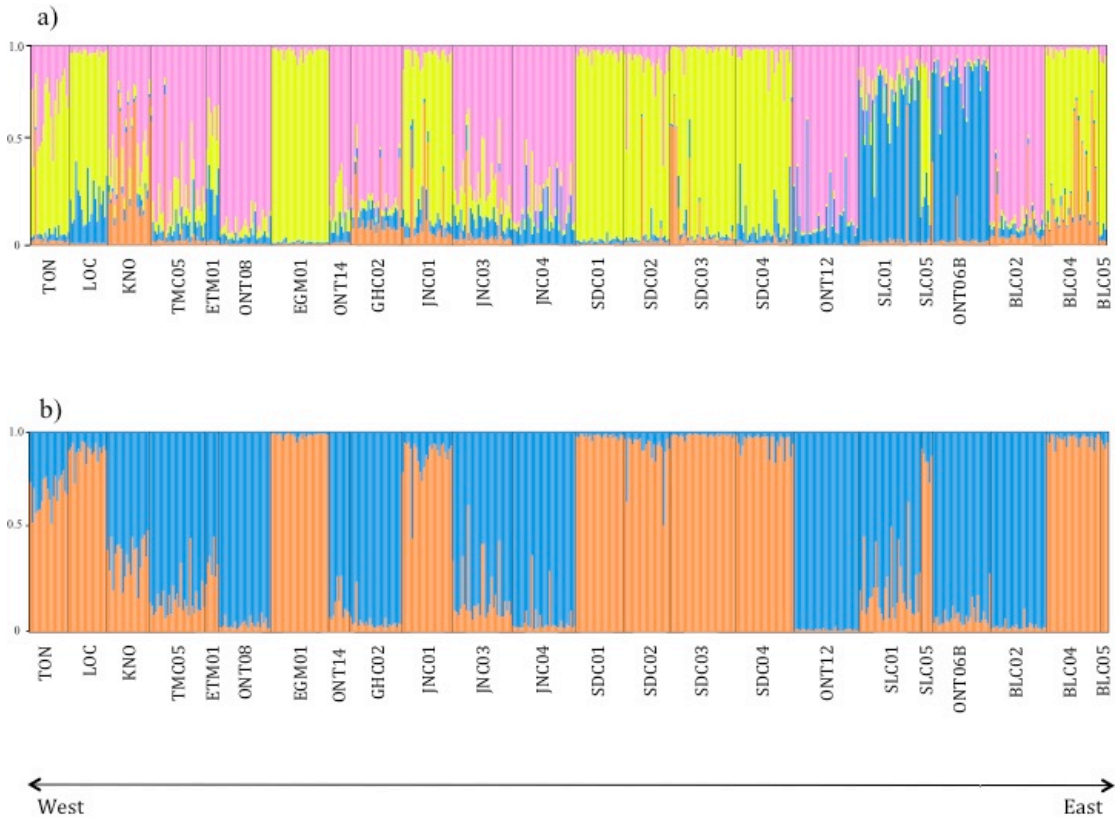
Appendix 2.5 cont'

Occupancy	Detection	AIC	ΔAIC	AIC	w_i	No. Parameter	Model Likelihood
B_17_dist,subs,distXspecies	(Uncond),subsXspecies(Uncond)	176.44	31.9	0	0	9	158.44
B_16_subs,subsXspeceis	(Uncond),	177.21	32.67	0	0	7	163.21
B_02_dist,distXspecies	(Uncond),_C	177.42	32.88	0	0	8	161.42
A_33_subs,	Constant	177.66	33.12	0	0	7	163.66
A_01_dist,	Constant	178.73	34.19	0	0	7	164.73
B_01_dist,_C	Constant	178.73	34.19	0	0	7	164.73
A_psi(.)_	Constant	179.53	34.99	0	0	6	167.53
B_psi(.)_C	Constant	179.78	35.24	0	0	6	167.78
B_02_dist,distXspecies	(Uncond),	181.49	36.95	0	0	7	167.49
B_09_dist,subs,_C	Constant	181.85	37.31	0	0	8	165.85
B_01_dist,	Constant	182.05	37.51	0	0	6	170.05
B_psi(.)_	Constant	182.89	38.35	0	0	5	172.89
B_10_dist,subs,distXspecies	(Uncond),_C	183.11	38.57	0	0	9	165.11
B_15_subs,_C	Constant	183.21	38.67	0	0	7	169.21
B_09_dist,subs,	Constant	186.39	41.85	0	0	7	172.39
B_15_subs,	Constant	187.16	42.62	0	0	6	175.16
B_10_dist,subs,distXspecies	(Uncond),	188.01	43.47	0	0	8	172.01
perfectdet_B_19_C_r	Constant	8355.57	8211.03	0	0	12	8331.57
perfectdet_A_15_C_r	Constant	8357.64	8213.1	0	0	12	8333.64

Appendix 2.7. List of species detected arranged by sites' link magnitude (✓, present; ✗, absent).

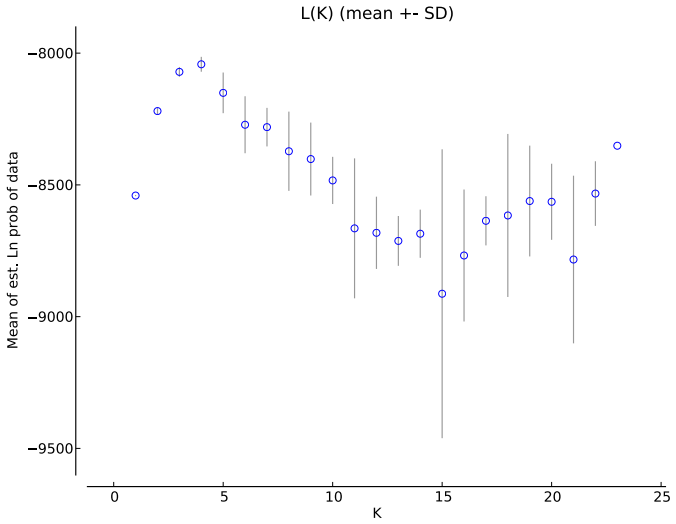
Latin name	Common name	Family	Link magnitude		
			4-8	9-15	16-51
Ambloplites rupestris	Rock bass	Centarchidae	✓	✓	✓
Ameiurus nebulosus	Brown bullhead	Ictaluridae	✗	✗	✓
Ameiurus sp.		Ictaluridae	✓	✗	✓
Amia calva	Bowfin	Amiidae	✗	✗	✓
Campostoma anomalum	Central stoneroller	Cyprinidae	✓	✓	✓
Campostoma sp.		Catostomidae	✓	✓	✓
Catostomus commersoni	White sucker	Catostomidae	✓	✓	✓
Culaea inconstans	Brook stickleback	Gasterosteidae	✓	✗	✗
Cyprinella sp.		Cyprinidae	✗	✓	✓
Cyprinus carpio	Common carp	Cyprinidae	✗	✓	✓
Esox lucius	Northern pike	Esocidae	✓	✓	✓
Etheostoma blennioides	Greenside darter	Percidae	✓	✗	✗
Etheostoma caeruleum	Rainbow darter	Percidae	✓	✗	✓
Etheostoma exile	Iowa darter	Percidae	✓	✗	✗
Etheostoma flabellare	Fantail darter	Percidae	✓	✗	✓
Etheostoma olmstedi	Teesellated darter	Percidae	✓	✗	✓
Fundulus diaphanus	Banded killifish	Fundulidae	✓	✓	✓
Hypentelium nigricans	Northern hogsucker	Catostomidae	✓	✓	✓
Ictalurus punctatus	Channel catfish	Ictaluridae	✗	✗	✓
Lepisosteus sp.	Generic gar	Lepisosteidae	✗	✓	✗
Lepomis cyanellus	Green sunfish	Centarchidae	✓	✗	✓
Lepomis gibbosus	Pumpkinseed sunfish	Centarchidae	✓	✓	✓
Lepomis macrochirus	Bluegill	Centarchidae	✓	✓	✓
Lepomis sp.		Centarchidae	✓	✗	✗
Luxilus cornutus	Common shiner	Cyprinidae	✓	✗	✓
Luxilus sp.		Cyprinidae	✓	✓	✓
Micropterus dolomieu	Smallmouth bass	Centarchidae	✗	✗	✓
Micropterus salmodies	Largemouth bass	Centarchidae	✓	✓	✓
Moxostoma sp.		Catostomidae	✓	✓	✓
Neogobius melanostomus	Round goby	Gobiidae	✓	✓	✓
Nocomis biguttatus	Hornyhead chub	Cyprinidae	✓	✗	✗
Nocomis micropogon	River chub	Cyprinidae	✓	✓	✓
Nocomis sp.		Cyprinidae	✓	✓	✓
Notemigonus crysoleucas	Golden shiner	Cyprinidae	✓	✓	✓
Notropis hudsonius	Spottail shiner	Cyprinidae	✗	✗	✓
Notropis sp.	Shiner	Cyprinidae	✓	✓	✓
Noturus flavus	Stonecat	Ictaluridae	✓	✗	✓
Noturus gyrinus	Tadpole madtom	Ictaluridae	✓	✗	✗
Noturus miurus	Brindled madtom	Ictaluridae	✗	✓	✗
Noturus sp.		Ictaluridae	✗	✗	✓
Oncorhynchus mykiss	Rainbow trout	Salmonidae	✓	✗	✗
Perca flavecens	Yellow perch	Percidae	✓	✓	✓
Percina caprodes	Logperch	Percidae	✓	✓	✓
Percina maculata	Blackside darter	Percidae	✓	✓	✓
Pimephales notatus	Bluntnose minnow	Cyprinidae	✓	✓	✓
Pimephales sp.		Cyprinidae	✓	✗	✗
Pimephales vigilax	Bullhead minnow	Cyprinidae	✓	✗	✗
Semotilus atromaculatus	Creek chub	Cyprinidae	✓	✗	✓
Umbra limi	Central mudminnow	Umbridae	✓	✗	✓
		Sampled sites	15	8	16
		Species	38	23	38

Appendix 3.1. (a) Based on a plot of estimated log likelihood for the data, the best model was $K = 4$, while a plot of the ΔK statistics of (Evanno et al. 2005) indicated $K = 2$ (Appendix 3.2). By implementing LnP(K=4) and the 0.6 threshold, all individuals from site ETM01 will be unassigned to any clusters and sites and clusters association were biologically hard to interpret (Appendix 3). (b) $K = 2$ based on DK statistics was determined to be the best model because clusters were interpretably associated with sites and the $K = 2$ had the least variance (see Appendix 4).

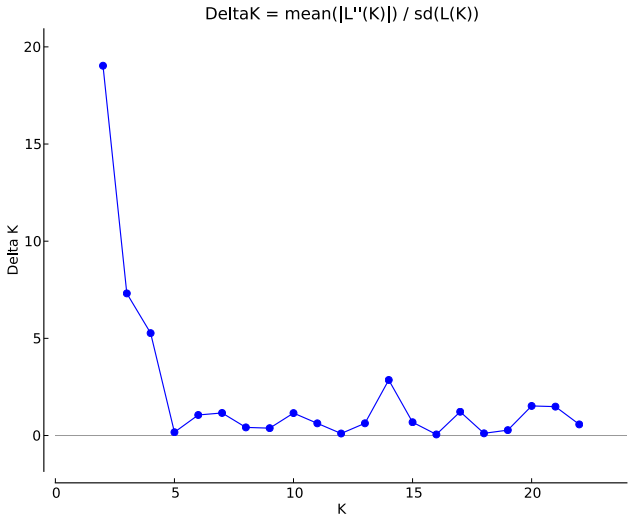


Appendix 3.2. a) Log probability of the data [LnP (D)] against K for five independent runs at each value (means and standard deviations). Note that some values of K>2 have a lower likelihood as K=2 but exhibit greater variance. Higher values of K also did not reveal any new genetic clusters consisting of individuals with high posterior probabilities of single group membership. Note also the at *K*=23, there was no SD was calculated as simulation was halted prematurely since LnP(D) was relatively stable b) ΔK (calculated according to Evanno et al., 2005) against K, illustrating the greatest change in likelihood at K= 2.

a)



b)



REFERENCES

Akaike H (1978) A Bayesian analysis of the minimum AIC procedure. Annals of the Institute of Statistical Mathematics 30:9–14

Balshine S, Verma A, Chant V and Theysmeyer T (2005) Competitive interactions between round gobies and logperch. Journal of Great Lakes Research 31:68–77

Bartels P, Hirsch PE, Svanbäck R and Eklöv P (2012) Water transparency drives intra-population divergence in Eurasian Perch (*Perca fluviatilis*). PloS one 7:e43641

Beermann J and Franke H-D (2012) Differences in resource utilization and behaviour between coexisting Jassa species (Crustacea, Amphipoda). Marine Biology 159:951–957

Beneteau CL, Walter RP, Mandrak NE and Heath DD (2011) Range expansion by invasion: genetic characterization of invasion of the greenside darter (*Etheostoma blennioides*) at the northern edge of its distribution. Biological Invasions 14:191–201

Bergstrom MA, Evrard LM and Mensinger AF (2008) Distribution, abundance, and range of the round goby, *Apollina melanostoma*, in the Duluth-Superior Harbor and St . Louis River Estuary, 1998 – 2004. Journal of Great Lakes Research 34:535–543

Bergstrom MA and Mensinger AF (2009) Interspecific resource competition between the invasive round goby and three native species: logperch, slimy sculpin, and spoonhead sculpin. Transactions of the American Fisheries Society 138:1009–1017

Bhagat Y, Fox MG and Ferreira MT (2011) Morphological diversification in introduced pumpkinseed (*Lepomis gibbosus*): assessing truss-based and geometric morphometric approaches. Fundamental and Applied Limnology 178:341–351

Björklund M and Almqvist G (2010) Rapid spatial genetic differentiation in an invasive species, the round goby *Neogobius melanostomus* in the Baltic Sea. Biological Invasions 12:2609–2618

Brinsmead J and Fox MG (2002) Morphological variation between lake- and stream-dwelling rock bass and pumpkinseed populations. Journal of Fish Biology 61:1619–1638

Bronnenhuber JE, Dufour BA, Higgs DM and Heath DD (2011) Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round

goby, *Neogobius melanostomus*, in Great Lakes tributaries. Molecular Ecology 20:1845–1859

Brown JE and Stepien CA (2008) Ancient divisions, recent expansions: phylogeography and population genetics of the round goby *Apollonia melanostoma*. Molecular Ecology 17:2598–2615

Brown JE and Stepien CA (2009) Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. Molecular Ecology 18:64–79

Brownscombe JW and Fox MG (2012) Range expansion dynamics of the invasive round goby (*Neogobius melanostomus*) in a river system. Aquatic Ecology 46:175–189

Brush JM, Fisk AT, Hussey NE, Johnson TB (2006) Spatial and seasonal variability in the diet of round goby (*Neogobius melanostomus*): stable isotopes indicate that stomach contents overestimate the importance of dreissenids. Canadian Journal of Fisheries and Aquatic Sciences 69:573–586

Budney LA and Hall BK (2010) Comparative morphology and osteology of pelvic fin-derived midline suckers in lumpfishes, snailfishes and gobies. Journal of Applied Ichthyology 26:167–175

Burnham KP and Anderson DR (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection. Sociological Methods & Research 33:261–304

Byholm P, Burgas D, Virtanen T and Valkama J (2012) Competitive exclusion within the predator community influences the distribution of a threatened prey species. Ecology 93:1802–1808

Chen Q, Wang C, Lu G, Zhao J, Chapman DC, Zsigmond J and Li S (2012) Microsatellite genetic diversity and differentiation of native and introduced grass carp populations in three continents. Genetica

Clapp DF, Schneeberger PJ, Jude DJ, Madison G and Pistis C (2001) Monitoring round goby (*Neogobius melanostomus*) population expansion in Eastern and Northern Lake Michigan. Journal of Great Lakes Research 27:335–341

Clarke LM, Munch SB, Thorrold SR and Conover DO (2010) High connectivity among locally adapted populations of a marine fish (*Menidia menidia*). Ecology 91:3526–37

Connolly SR, Menge BA and Roughgarden J (2001) A latitudinal gradient in recruitment of intertidal invertebrates in the northeast Pacific Ocean. Ecology 82:1799–1813

Cooper MJ, Ruetz III CR, Uzarski DG and Burton TM (2007) Distribution of round gobies in coastal areas of Lake Michigan: Are wetlands resistant to invasion? Journal of Great Lakes Research 33:303–313

Corkum LD, MacInnis AJ and Wickett RG (1998) Reproductive habits of round gobies. Great Lakes Research Review 3:13–20

Cornuet JM and Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

Cosentino BJ, Schooley RL and Phillips CA (2011) Spatial connectivity moderates the effect of predatory fish on salamander metapopulation dynamics. Ecosphere 2:art95

Coulon A, Fitzpatrick JW, Bowman R, Stith BM, Makarewich CA, Stenzler LM and Lovette J (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). Molecular Ecology 17:1685–1701

Coulter DP, Murry BA and Uzarski DG (2012) Use of wetland versus open habitats by round gobies in lakes Michigan and Huron: Patterns of CPUE, length, and maturity. Journal of Great Lakes Research 38:439–444

Cucherousset J and Olden JD (2011) Ecological impacts of non-native freshwater fishes. Fisheries 36:215–230

Dettmers JM, Goddard CI and Smith KD (2012) Management of alewife using Pacific salmon in the Great Lakes : Whether to manage for economics or the ecosystem? Fisheries 37:37–41

Dillon AK and Stepien CA (2001) Genetic and biogeographic relationships of the invasive round (*Neogobius melanostomus*) and tubenose (*Proterorhinus marmoratus*) gobies in the Great Lakes versus Eurasian populations. Journal of Great Lakes Research 27:267–280

Dlugosch KM and Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology 17:431–449

Dubs DOL and Corkum LD (1996) Behavioral interactions between round gobies (Neogobius melanostomus) and mottled sculpins (*Cottus bairdi*). Journal of Great Lakes Research 22:838–844

Dufour B, Hogan TM and Heath DD (2007) Ten polymorphic microsatellite markers in the invasive round goby (*Neogobius melanostomus*) and cross-species amplification. Molecular Ecology Notes 7:1205–1207

Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620

Excoffier L and Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564–567

Falke JA, Bailey LL, Fausch KD and Bestgen KR (2012) Colonization and extinction in dynamic habitats: an occupancy approach for a Great Plains stream fish assemblage. Ecology 93:858–867

French III JRP and Jude DJ (2001) Diets and diet overlap of nonindigenous gobies and small benthic native fishes co-inhabiting the St. Clair River, Michigan. Journal of Great Lakes Research 27:300–311

García-Berthou E (2007) The characteristics of invasive fishes: what has been learned so far? Journal of Fish Biology 71:33–55

Gillespie GJ and Fox MG (2003) Morphological and life-history differentiation between littoral and pelagic forms of pumpkinseed. Journal of Fish Biology 62:1099–1115

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity 86:485–486

Gozlan RE (2008) Introduction of non-native freshwater fish: is it all bad? Fish and Fisheries 9:106–115

Grigorovich IA, Colautti RI, Mills EL, Holeck K, Ballert AG and MacIsaac HJ (2003) Ballast-mediated animal introductions in the Laurentian Great Lakes: retrospective and prospective analyses. Canadian Journal of Fisheries and Aquatic Sciences 756:740–756

Groce MC, Bailey LL and Fausch KD (2012) Evaluating the success of Arkansas darter translocations in Colorado: an occupancy sampling approach. Transactions of the American Fisheries Society 141:825–840

Gu W and Swihart RK (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biological Conservation 116:195–203

Gutowsky LFG and Fox MG (2012) Intra-population variability of life-history traits and growth during range expansion of the invasive round goby, *Neogobius melanostomus*. Fisheries Management and Ecology 19:78–88

Haas TC, Blum MJ and Heins DC (2010) Morphological responses of a stream fish to water impoundment. Biology letters 6:803–806

Hayden TA and Miner JG (2008) Rapid dispersal and establishment of a benthic Ponto-Caspian goby in Lake Erie: diel vertical migration of early juvenile round goby. Biological Invasions 11:1767–1776

Hendry AP, Day T and Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. Evolution 55:459–466

Hensler SR and Jude DJ (2007) Diel vertical migration of round goby larvae in the Great Lakes diel vertical migration of round goby larvae in the Great Lakes. Journal of Great Lakes Research 33:295–302

Hinch SG, Somers KM and Collins NC (1994) Autocorrelation and assessment of habitat-abundance relationships in littoral zone fish. Canadian Journal of Fisheries and Aquatic Sciences 51:701–712

Hines JE (2006) PRESENCE2- Software to estimate patch occupancy and related parameters. USGS-PWRC. http://www.mbr-pwrc.usgs.gov/software/presence.html

Holeck KT, Mills EL, Macisaac HJ, Dochoda MR, Colautti RI and Ricciardi A (2004) Bridging troubled waters: biological invasions, transoceanic shipping, and the Laurentian Great Lakes. BioScience 54:919–929

Horn PL, Forman JS and Dunn MR (2012) Dietary partitioning by two sympatric fish species, red cod (*Pseudophycis bachus*) and sea perch (*Helicolenus percoides*), on Chatham Rise , New Zealand. Marine Biology Research 8:624–634

Hubisz MJ, Falush D, Stephens M and Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322–1332

Jakobsson M and Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Janssen J and Jude DJ (2001) Recruitment failure of mottled sculpin Cottus bairdi in Calumet Harbor, Southern Lake Michigan, induced by the newly introduced round goby Neogobius melanostomus. Journal of Great Lakes Research 27:319–328

Johnson TB, Bunnell DB and Knight CT (2005) A potential new energy pathway in central Lake Erie: the round goby connection. Journal of Great Lakes Research 31:238–251

Jude DJ, Reider RH and Smith W (1992) Establishment of Gobiidae in the Great Lakes Basin. Canadian Journal of Fisheries and Aquatic Sciences 49:416–421

Keller RP, Lodge DM and Finnoff DC (2007) Risk assessment for invasive species produces net bioeconomic benefits. Proceedings of the National Academy of Sciences of the United States of America 104:203–207

Kipp R and Ricciardi A (2012) Impacts of the Eurasian round goby (*Neogobius melanostomus*) on benthic communities in the upper St . Lawrence River. Canadian Journal of Fisheries and Aquatic Sciences 69:469–486

Klima K, Travis SE (2012) Genetic population structure of invasive nutria (*Myocastor coypus*) in Louisiana, USA: is it sufficient for the development of eradication units? Biological Invasions 14:1909–1918

Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. Molecular Ecology Resources 11:353–357

Kornis M., Mercado-Silva N and Vander Zanden MJ (2012) Twenty years of invasion: a review of round goby *Neogobius melanostomus* biology, spread and ecological implications. Journal of Fish Biology 80:235–285

Kornis MS and Vander Zanden MJ (2010) Forecasting the distribution of the invasive round goby (Neogobius melanostomus) in Wisconsin tributaries to Lake Michigan. Canadian Journal of Fisheries and Aquatic Sciences 67:553–562

Krakowiak PJ and Pennuto CM (2008) Fish and macroinvertebrate communities in tributary streams of eastern Lake Erie with and without round gobies (*Neogobius melanostomus*, Pallas 1814). Journal of Great Lakes Research 34:675–689

Langerhans BR (2008) Predictability of phenotypic differentiation across flow regimes in fishes. Integrative and Comparative Biology 48:750–768

Langerhans RB, Layman CA and Langerhans AK, Dewitt TJ (2003) Habitat-associated morphological divergence in two Neotropical fish species. Biological Journal of the Linnean Society 80:689–698.

Langerhans RB and Reznick DN (2010) Ecology and evolution of swimming performance in fishes: Predicting evolution with biomechanics. In: Domenici P, Kapoor BG (eds) Fish locomotion: An eco-ethological perspective. NH: Science Publishers, Enfield, pp 200–248

LaRue EA, Ruetz CR, Stacey MB and Thum R (2011) Population genetic structure of the round goby in Lake Michigan: implications for dispersal of invasive species. Hydrobiologia 663:71–82

Lauder GV and Drucker EG (2004) Morphology and experimental hydrodynamics of fish fin control surfaces. IEEE Journal of Oceanic Engineering 29:556–571

Lee CE (2002) Evolutionary genetics of invasive species. Trends in Ecology and Evolution 17:386–391

Luikart G, Sherwin WB, Steele BM and Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology 7:963–974

Lynch MP and Mensinger AF (2012) Seasonal abundance and movement of the invasive round goby (*Neogobius melanostomus*) on rocky substrate in the Duluth-Superior Harbor of Lake Superior. Ecology of Freshwater Fish 21:64–74

L'avrincikova M, Kovac V and Katina S (2005) Ontogenetic variability in external morphology of round goby *Neogobius melanostomus* from Middle Danube, Slovakia. Journal of Applied Ichthyology 21:328–334

MacInnis AJ and Corkum LD (2000a) Fecundity and reproductive season of the round goby *Neogobius melanostomus* in the Upper Detroit River. Transactions of the American Fisheries Society 129:852–858

MacInnis AJ and Corkum LD (2000b) Age and growth of round goby *Neogobius melanostomus* in the Upper Detroit River. Transactions of the American Fisheries Society 129:852–858

Mackenzie DI, Bailey LL and Nichols JD (2004) Investigating species co-occurrence patterns when species. Journal of Animal Ecology 73:546–555

MacKenzie DI, Nichols JD, Lachman GB, Droege, S, Royle, JA and Langtimm CA (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255

Madenjian CP, Stapanian M a., Witzel LD, et al. (2011) Evidence for predatory control of the invasive round goby. Biological Invasions 13:987–1002. doi: 10.1007/s10530-010-9884-7

Marentette JR, Tong S, Wang G, Sopinka NM, Taves MD, Koops MA and Balshine S (2012) Behavior as biomarker? Laboratory versus field movement in round goby (*Neogobius melanostomus*) from highly contaminated habitats. Ecotoxicology 21:1003–1012

Marsden JE, Charlebois P and Wolfe K (1996) The round goby (*Neogobius melanostomus*): a review of European and North American literature. Zion, Illinois

Meixler MS, Arend KK and Bain MB (2005) Fish community support in wetlands within protected embayments of Lake Ontario. Journal of Great Lakes Research 31:188–196

Meixler MS and Bain MB (2010) Landscape scale assessment of stream channel and riparian habitat restoration needs. Landscape and Ecological Engineering 6:235–245

Mills MD, Rader RB and Belk MC (2004) Complex interactions between native and invasive fish: the simultaneous effects of multiple negative interactions. Oecologia 141:713–21

Mittelbach GG, Osenberg CW and Wainwright PC (1999) Variation in feeding morphology between pumpkinseed populations: Phenotypic plasticity or evolution? Evolutionary Ecology Research 1:111–128

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Maruyama T and Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29:1–10

Nett JHG, Campbell TB, Mandrak NE and Tiegs SD (2012) Detecting invasive round goby in wadeable streams: A comparison of gear types. North American Journal of Fisheries Management 32:360–364

Nolte AW (2011) Dispersal in the course of an invasion. Molecular Ecology 20:1803–1804

Olden JD, Poff LN and Bestgen KR (2006) Life-history strategies predict fish invasions and extirpations in the Colorado River Basin. Ecological Monographs 76:25–40

Van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P (2004) Micro-Checker: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535–538

Paetkau D, Slade R, Burden M and Arnaud E (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13:55–65

Page LM and Schemske DW (1978) The effect of interspecific competition on the distribution and size of darters of the subgenus Catonotus (Percidae: Etheostoma). Copeia 3:406–412

Paul MJ and Meyer JL (2001) Streams in the urban landscape. Annual Review of Ecology and Systematics 32:333–365

Pennuto CM, Krakowiak PJ and Janik CE (2010) Seasonal abundance, diet, and energy consumption of round gobies (*Neogobius melanostomus*) in Lake Erie tributary streams. Ecology of Freshwater Fish 19:206–215

Peres-Neto PR (2004) Patterns in the co-occurrence of fish species in streams: the role of site suitability, morphology and phylogeny versus species interactions. Oecologia 140:352–360

Phillips EC, Washek ME, Hertel AW and Niebel BM (2003) The round goby (*Neogobius melanostomus*) in Pennsylvania tributary streams of Lake Erie. Journal of Great Lakes Research 29:34–40

Pilliod DS, Arkle RS and Maxell BA (2012) Persistence and extirpation in invaded landscapes: patch characteristics and connectivity determine effects of non-native predatory fish on native salamanders. Biological Invasions. doi: 10.1007/s10530-012-0317-7

Pimentel D, Zuniga R and Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 52:273–288

Piry S, Alapetite A, Cornuet J-M, Baudduin, L and Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95:536–539

Poos M, Dextrase AJ, Schwalb AN and Ackerman JD (2009) Secondary invasion of the round goby into high diversity Great Lakes tributaries and species at risk hotspots: potential new concerns for endangered freshwater species. Biological Invasions 12:1269–1284

Poulos HM, Chernoff B, Fuller PL and Butman D (2012) Ensemble forecasting of potential habitat for three invasive fishes. Aquatic Invasions. doi: 10.3391/ai.2012.7.01

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pritchard JK, Wen X and Falush D (2007) Documentation for STRUCTURE software: Version 2.2

Raby GD, Gutowsky LFG, Fox MG (2010) Diet composition and consumption rate in round goby (*Neogobius melanostomus*) in its expansion phase in the Trent River, Ontario. Environmental Biology of Fishes 89:143–150

Rahel FJ and Olden JD (2008) Assessing the effects of climate change on aquatic invasive species. Conservation Biology 22:521–533

Ralph AL, Berli BI, Burkhardt-Holm P and Tierney KB (2012) Variability in swimming performance and underlying physiology in rainbow trout (*Oncorhynchus mykiss*) and brown trout (*Salmo trutta*). Comparative Biochemistry and Physiology Part A, Molecular & Integrative Physiology 163:350–356

Randall RG, Kelso JRM and Minns CK (1995) Fish production in freshwaters: Are rivers more productive than lakes? Canadian Journal of Fisheries and Aquatic Sciences 51:631–643

Randall RG, Minns CK, Cairns VW and Moore JE (1996) The relationship between an index of fish production and submerged macrophytes and other habitat features at three littoral areas in the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 53:35–44

Rannala B and Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the United States of America 94:9197–9201

Ray WJ and Corkum LD (2001) Habitat and site affinity of the round goby. Journal of Great Lakes Research 27:329–334

Raymond M and Rousset F (1995) GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Richmond OMW, Hines JE and Beissinger SR (2010) Two-species occupancy models: a new parameterization applied to co-occurrence of secretive rails. Ecological Applications 20:2036–2046

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M and Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations.

Proceedings of the National Academy of Sciences of the United States of America 91:3166–3170

Rixon CAM, Duggan IC, Bergeron NMN, Ricciardi A and MacIsaac HJ (2005) Invasion risks posed by the aquarium trade and live fish markets on the Laurentian Great Lakes. Biodiversity and Conservation 14:1365–1381

Roman J and Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. Trends in Ecology and Evolution 22:454–464

Rosenberg NA (2004) Distruct: A program for the graphical display of population structure. Molecular Ecology Notes 4:137–138

Sanford E and Worth DJ (2010) Local adaptation along a continuous coastline: prey recruitment drives differentiation in a predatory snail. Ecology 91:891–901

Schlosser IJ and Toth LA (1984) Niche relationships and population ecology of rainbow ecology (*Etheostoma caeruleum*) and fantail (*E. flabellare*) darters in a temporally variable environment. Oikos 42:229–238

Selkoe KA and Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters 9:615–629

Shreve RRL (1966) Statistical law of stream numbers. Journal of Geology 74:17–37

Simonovicå P, Paunovicå M and Popovicå S (2001) Morphology, feeding and reproduction of the round goby , *Neogobius melanostomus* (Pallas), in the Danube River Basin, Yugoslavia. Journal of Great Lakes Research 27:281–289

Singkran N and Meixler MS (2008) Influences of habitat and land cover on fish distributions along a tributary to Lake Ontario, New York. Landscape Ecology 23:539–551

Smith SA and Donoghue MJ (2008) Rates of molecular evolution are linked to life history in flowering plants. Science 322:86–89

Van Snik Gray E, Kellogg KA and Stauffer JR (2005) Habitat shift of a native darter *Etheostoma olmstedi* (Teleostei: Percidae) in sympatry with a non-native darter *Etheostoma zonale*. American Midland Naturalist 154:166–177

Snik Gray E Van, Stauffer JR (2001) Substrate choice by three species of darters (Teleostei: Percidae) in an artificial stream: effects of a nonnative species. Copeia 1:254–261

Stanley TR and Royle JA (2005) Estimating site occupancy and abundance using indirect detection indices. Journal of Wildlife Management 69:874–883

Stepien CA, Brown JE, Neilson ME and Tumeo MA (2005) Genetic diversity of invasive species in the Great Lakes versus their Eurasian source populations: insights for risk analysis. Risk Analysis 25:1043–1060

Stepien CA and Tumeo MA (2006) Invasion genetics of Ponto-Caspian gobies in the Great Lakes: a "cryptic" species, absence of founder effects, and comparative risk analysis. Biological Invasions 8:61–78

Sterling KA, Reed DH, Noonan BP and Warren ML (2012) Genetic effects of habitat fragmentation and population isolation on *Etheostoma raneyi* (Percidae). Conservation Genetics 13:859–872

Swain DP and Holtby LB (1989) Differences in morphology and behavior between juvenile coho salmon (*Oncorhynchus kisutch*) rearing in a Lake and in its tributary stream. Canadian Journal of Fisheries and Aquatic Sciences 46:1406–1414

Taraborelli AC, Fox MG, Schaner T and Johnson TB (2009) Density and habitat use by the round goby (*Apollonia melanostoma*) in the Bay of Quinte, Lake Ontario. Journal of Great Lakes Research 35:266–271

Tierney KB, Kasurak A V., Zielinski BS and Higgs DM (2011) Swimming performance and invasion potential of the round goby. Environmental Biology of Fishes 92:491–502

Vila-Gispert A, Fox MG, Zamora L and Moreno-Amich R (2007) Morphological variation in pumpkinseed *Lepomis gibbosus* introduced into Iberian lakes and reservoirs; adaptations to habitat type and diet? Journal of Fish Biology 71:163–181

Viscosi V and Cardini A (2011) Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. PloS ONE 6:e25630

Vélez-Espino LA, Koops MA and Balshine S (2010) Invasion dynamics of round goby (*Neogobius melanostomus*) in Hamilton Harbour, Lake Ontario. Biological Invasions 12:3861–3875

Vøllestad LA, Varreng K and Poleo ABS (2004) Body depth variation in crucian carp Carassius carassius: an experimental individual-based study. Ecology of Freshwater Fish 13:197–202

Wagner CE and McCune AR (2009) Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. Evolution 63:1312–1326

Walsh MG, Dittman DE and Gorman RO (2007) Occurrence and food habits of the round goby in the profundal zone of southwestern Lake Ontario. Journal of Great Lakes Research 33:83–92

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity 89:438–450

Wei A, Chow-Fraser P and Albert D (2004) Influence of shoreline features on fish distribution in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 61:1113–1123

Westphal MI, Browne M, MacKinnon K and Noble I (2007) The link between international trade and the global distribution of invasive alien species. Biological Invasions 10:391–398

Williams BL, Brawn JD and Paige KN (2003) Landscape scale genetic effects of habitat fragmentation on a high gene flow species: *Speyeria idalia* (Nymphalidae). Molecular Ecology 12:11–20

Wimberger PH (1992) Plasticity of fish body shape. The effects of diet , development , family and age in two species of Geophugus (Pisces: Cichlidae). Biological Journal of the Linnean Society 45:197–218

Wolfe KR and Marsden EJ (1998) Tagging methods for the round goby (*Neogobius melanostomus*). Journal of Great Lakes Research 24:731–735

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Wright S (1943) Isolation by distance. Genetics 28:114–137

Yavno S, Fox MG, Vila-Gispert A and Bhagat Y (2012) Morphological differences between native and non-native pumpkinseed in traits associated with locomotion. Environmental Biology of Fishes. doi: 10.1007/s10641-012-0053-y

Young JM, Marentette JR, Gross C, McDonald JI, Verma A, Marsh-Rollo SE, Mcdonald PDM, Earn DJD and Balshine S (2010) Demography and substrate affinity of the round goby (*Neogobius melanostomus*) in Hamilton Harbour. Journal of Great Lakes Research 36:115–122

Vander Zanden MJ and Olden JD (2008) A management framework for preventing the secondary spread of aquatic invasive species. Canadian Journal of Fisheries and Aquatic Sciences 65:1512–1522

Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2004) Geometric morphometrics for biologists: A primer. CA Academic Press. San Diego

Zhan A, Darling JA, Bock DG, Lacoursiere-Roussel, A, MacIsaac, HJ and Cristescu ME (2012) Complex genetic patterns in closely related colonizing invasive species. Ecology and Evolution 2:1331–1346